

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 5, 2004, 14:09:33 ; Search time 54 Seconds  
(without alignments)  
544.166 Million cell updates/sec

Title: US-09-997-428-408  
Perfect score: 502  
Sequence: 1 MKIAALLGLCVLSCSSAAA.....QAVGAVKALKALCALTVFG 104

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

- Database : A\_Geneseq\_29Jan04: \*
- 1: geneseqp1980s: \*
  - 2: geneseqp1990s: \*
  - 3: geneseqp2000s: \*
  - 4: geneseqp2001s: \*
  - 5: geneseqp2002s: \*
  - 6: geneseqp2003as: \*
  - 7: geneseqp2003bs: \*
  - 8: geneseqp2004s: \*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	502	100.0	104	3	AAY66757 Membrane-
2	502	100.0	104	3	AAY44458 Human lun
3	502	100.0	104	3	AAY87288 Human sig
4	502	100.0	104	4	AAB65280 Human PRO
5	502	100.0	104	5	AAB61141 Human PRO
6	502	100.0	104	6	ABU58095 Human PRO
7	502	100.0	104	6	ABU59173 Novel hum
8	502	100.0	104	6	ABU82865 Human sec
9	502	100.0	104	6	AAO19895 Human ute
10	502	100.0	104	6	ABU60604 Human sec
11	502	100.0	104	6	ABU13986 Human PRO
12	502	100.0	104	6	ABU72571 Novel hum
13	502	100.0	104	6	ABU59320 Human sec
14	502	100.0	104	6	ABO26017 Human PRO
15	502	100.0	104	6	ABU59026 Human sec
16	502	100.0	104	6	ABU92404 Novel hum
17	502	100.0	104	6	ABU59469 Human hum
18	502	100.0	104	6	ABU92235 Novel hum
19	502	100.0	104	6	ABU10941 Human PRO
20	502	100.0	104	6	ABU81693 Novel hum
21	502	100.0	104	6	ABU88632 Human sec
22	502	100.0	104	6	ABO34146 Human PRO
23	502	100.0	104	6	ADA37919 Human sec
24	502	100.0	104	6	ADA21605 Human sec
25	502	100.0	104	6	ADA10392 Human sec

26	502	100.0	104	6	ADA17936	Human PRO
27	502	100.0	104	6	ADA28044	Human sec
28	502	100.0	104	6	ADA94624	Human sec
29	502	100.0	104	6	ADA38849	Human sec
30	502	100.0	104	6	ADA92970	Human sec
31	502	100.0	104	7	ABO53232	Human sec
32	502	100.0	104	7	ADA22531	Human sec
33	502	100.0	104	7	ABO22602	Human sec
34	502	100.0	104	7	ADA06697	Human sec
35	502	100.0	104	7	ADA39390	Human sec
36	502	100.0	104	7	ADC96416	Human PRO
37	502	100.0	104	7	ADC57888	Human PRO
38	502	100.0	104	7	ADC55252	Human PRO
39	502	100.0	104	7	ADC12119	Human sec
40	502	100.0	104	7	ADC58541	Human PRO
41	502	100.0	104	7	ADC07596	Human sec
42	502	100.0	104	7	ADC11586	Human sec
43	502	100.0	104	7	ADC14708	Novel hum
44	502	100.0	104	7	ADD08240	Novel hum
45	502	100.0	104	7	ADC82065	Human PRO
46	502	100.0	104	7	ADD07707	Novel hum
47	502	100.0	104	7	ADC82598	Human PRO
48	502	100.0	104	7	ADD08778	Novel hum
49	502	100.0	104	7	ADD07027	Novel hum
50	502	100.0	104	7	ADC83274	Human PRO
51	502	100.0	104	7	ADD55381	Human PRO
52	502	100.0	104	7	ADD56339	Human PRO
53	502	100.0	104	7	ADD54777	Human PRO
54	502	100.0	104	7	ADE26931	Novel hum
55	502	100.0	104	7	ADE26398	Novel hum
56	498	99.2	104	2	AAW75868	LUI05 pol
57	250	49.8	104	6	AAO19893	Murine ut
58	160	31.9	93	2	AAW62068	Human lun
59	160	31.9	93	2	AAW28334	LUI03 spe
60	160	31.9	93	3	AAW44456	Human lun
61	160	31.9	93	3	AAW87289	Human sig
62	160	31.9	93	4	AAU29145	Human PRO
63	160	31.9	93	6	ABU58521	Human PRO
64	160	31.9	93	6	ABU88069	Novel hum
65	160	31.9	93	6	ABU84384	Human sec
66	160	31.9	93	6	ABR66258	Human sec
67	160	31.9	93	6	ABR65648	Human sec
68	160	31.9	93	6	ABU93588	Human sec
69	160	31.9	93	6	ABU84827	Human PRO
70	160	31.9	93	6	ABU89948	Novel hum
71	160	31.9	93	6	ABR68197	Human sec
72	160	31.9	93	6	AAO19894	Human ute
73	160	31.9	93	6	ABU96250	Novel hum
74	160	31.9	93	6	ABU92681	Human sec
75	160	31.9	93	6	ABO08758	Human sec
76	160	31.9	93	6	ABO02810	Human sec
77	160	31.9	93	6	ABR74964	Human sec
78	160	31.9	93	6	ABR94726	Human sec
79	160	31.9	93	6	ABU85699	Human PRO
80	160	31.9	93	6	ABU98859	Novel hum
81	160	31.9	93	6	ABU98074	Novel hum
82	160	31.9	93	6	ABU93780	Human PRO
83	160	31.9	93	6	ABU89473	Human PRO
84	160	31.9	93	6	ABU86314	Human sec
85	160	31.9	93	6	ABU67527	Human sec
86	160	31.9	93	6	ABU80555	Human PRO
87	160	31.9	93	6	ABR99473	Human sec
88	160	31.9	93	6	ABR98863	Human sec
89	160	31.9	93	6	ABO16386	Human sec
90	160	31.9	93	6	ABR92286	Human sec
91	160	31.9	93	6	ABO18927	Human sec
92	160	31.9	93	6	ABR78348	Human sec
93	160	31.9	93	6	ABU85084	Novel hum
94	160	31.9	93	6	ABO00223	Novel hum
95	160	31.9	93	6	ABO11555	Human sec
96	160	31.9	93	6	ABO02200	Human sec
97	160	31.9	93	6	ABU88774	Novel hum
98	160	31.9	93	6	ABU83469	Human sec

99 160 31.9 93 6 ABO06270 Novel hum  
100 160 31.9 93 6 ABR59306 Abr59306 Human sec

## ALIGNMENTS

RESULT 1

AAY66757

ID AAY66757 standard; protein; 104 AA.

XX AC AAY66757;

XX DT 05-APR-2000 (first entry)

XX DE Membrane-bound protein PRO1245.

XX KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;  
XX KW pharmaceutical; receptor immunoadhesin; gene mapping.

XX OS Homo sapiens.

XX PN WO9963088-A2.

XX PD 09-DEC-1999.

XX PF 02-JUN-1999; 99WO-US012252.

XX PR 02-JUN-1998; 98US-0087607P.

XX PR 02-JUN-1998; 98US-0087609P.

XX PR 03-JUN-1998; 98US-0087759P.

XX PR 04-JUN-1998; 98US-0088021P.

XX PR 04-JUN-1998; 98US-0088025P.

XX PR 04-JUN-1998; 98US-0088028P.

XX PR 04-JUN-1998; 98US-0088029P.

XX PR 04-JUN-1998; 98US-0088030P.

XX PR 04-JUN-1998; 98US-0088033P.

XX PR 05-JUN-1998; 98US-0088167P.

XX PR 05-JUN-1998; 98US-0088202P.

XX PR 05-JUN-1998; 98US-0088212P.

XX PR 05-JUN-1998; 98US-0088217P.

XX PR 09-JUN-1998; 98US-0088655P.

XX PR 10-JUN-1998; 98US-0088722P.

XX PR 10-JUN-1998; 98US-0088730P.

XX PR 10-JUN-1998; 98US-0088734P.

XX PR 10-JUN-1998; 98US-0088738P.

XX PR 10-JUN-1998; 98US-0088740P.

XX PR 10-JUN-1998; 98US-0088741P.

XX PR 10-JUN-1998; 98US-0088742P.

XX PR 10-JUN-1998; 98US-0088810P.

XX PR 10-JUN-1998; 98US-0088811P.

XX PR 10-JUN-1998; 98US-0088824P.

XX PR 10-JUN-1998; 98US-0088825P.

XX PR 11-JUN-1998; 98US-0088858P.

XX PR 11-JUN-1998; 98US-0088861P.

XX PR 11-JUN-1998; 98US-0088863P.

XX PR 12-JUN-1998; 98US-0089090P.

XX PR 12-JUN-1998; 98US-0089105P.

XX PR 16-JUN-1998; 98US-0089440P.

XX PR 16-JUN-1998; 98US-0089512P.

XX PR 16-JUN-1998; 98US-0089514P.

XX PR 17-JUN-1998; 98US-0089532P.

XX PR 17-JUN-1998; 98US-0089538P.

XX PR 17-JUN-1998; 98US-0089598P.

XX PR 17-JUN-1998; 98US-0089599P.

XX PR 17-JUN-1998; 98US-0089600P.

XX PR 17-JUN-1998; 98US-0089653P.

XX PR 18-JUN-1998; 98US-0089801P.

XX PR 18-JUN-1998; 98US-0089907P.

PR 18-JUN-1998; 98US-0099908P.  
PR 19-JUN-1998; 98US-0099947P.  
PR 19-JUN-1998; 98US-0099948P.  
PR 19-JUN-1998; 98US-0099952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 23-JUN-1998; 98US-0090349P.  
PR 23-JUN-1998; 98US-0090355P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090431P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090445P.  
PR 24-JUN-1998; 98US-0090461P.  
PR 24-JUN-1998; 98US-0090472P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090538P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 24-JUN-1998; 98US-0090557P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090691P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 01-JUL-1998; 98US-0091358P.  
PR 01-JUL-1998; 98US-0091360P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091519P.  
PR 02-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091633P.  
PR 02-JUL-1998; 98US-0091646P.  
PR 02-JUL-1998; 98US-0091673P.  
PR 07-JUL-1998; 98US-0091978P.  
PR 07-JUL-1998; 98US-0091982P.  
PR 09-JUL-1998; 98US-0092182P.  
PR 10-JUL-1998; 98US-0092472P.  
PR 20-JUL-1998; 98US-0093339P.  
PR 30-JUL-1998; 98US-0094651P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 04-AUG-1998; 98US-0095285P.  
PR 04-AUG-1998; 98US-0095301P.  
PR 04-AUG-1998; 98US-0095302P.  
PR 04-AUG-1998; 98US-0095318P.  
PR 04-AUG-1998; 98US-0095321P.  
PR 04-AUG-1998; 98US-0095325P.  
PR 10-AUG-1998; 98US-0095916P.  
PR 10-AUG-1998; 98US-0095929P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 11-AUG-1998; 98US-0096143P.  
PR 11-AUG-1998; 98US-0096146P.  
PR 12-AUG-1998; 98US-0096329P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096768P.  
PR 17-AUG-1998; 98US-0096773P.  
PR 17-AUG-1998; 98US-0096791P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096894P.  
PR 17-AUG-1998; 98US-0096895P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096899P.  
PR 18-AUG-1998; 98US-0096950P.  
PR 18-AUG-1998; 98US-0096959P.

```

PR 18-AUG-1998; 98US-0096960P.
PR 18-AUG-1998; 98US-0097022P.
PR 19-AUG-1998; 98US-0097341P.
PR 20-AUG-1998; 98US-0097218P.
PR 24-AUG-1998; 98US-0097661P.
PR 26-AUG-1998; 98US-0097951P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0097978P.
PR 26-AUG-1998; 98US-0097979P.
PR 26-AUG-1998; 98US-0097986P.
PR 26-AUG-1998; 98US-0098014P.
PR 31-AUG-1998; 98US-0098525P.
PR 16-SEP-1998; 98US-0100634P.
PR 12-JAN-1999; 99US-0115565P.
XX
XX (GETH ) GENENTECH INC.
XX
XX Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
XX Wood WI, Yuan J;
XX
XX WPI; 2000-072883/06.
XX N-PSDB; AAZ65103.
XX
XX Membrane-bound proteins and related nucleotide sequences.
XX
XX Claim 12; Fig 290; 822pp; English.
XX
XX The invention provides membrane-bound PRO polypeptides and
XX polynucleotides encoding them. The PRO sequences of the invention were
XX identified based on extracellular domain homology screening. The PRO
XX sequences have homology with proteins including LDL receptors, TIE
XX ligands and various enzymes. The membrane-bound proteins and receptor
XX molecules are useful as pharmaceutical and diagnostic agents. Receptor
XX immunoadhesins, for instance, can be used as therapeutic agents to block
XX receptor-ligand interactions. The membrane-bound proteins can also be
XX employed for screening of potential peptide or small molecule inhibitors
XX of the relevant receptor/ligand interaction. The PRO encoding sequences
XX are useful as hybridization probes, in chromosome and gene mapping and in
XX the generation of antisense RNA and DNA. PRO nucleic acid sequences will
XX also be useful for the preparation of PRO polypeptides, especially by
XX recombinant techniques
XX
XX Sequence 104 AA;
XX
XX Query Match 100.0%; Score 502; DB 3; Length 104;
XX Best Local Similarity 100.0%; Pred. No. 1.4e-48;
XX Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MKLAALGLCVALSCHSSAAAFVLSGSAKPVAQPVAALESAAEAGAGTLANPLGTLNPKLL 60
XX DB 1 MKLAALGLCVALSCHSSAAAFVLSGSAKPVAQPVAALESAAEAGAGTLANPLGTLNPKLL 60
XX
XX QY 61 LSSLGIPVNHLEIGSKQVAEIGPQAVGAVKALKALLGALTIVFG 104
XX DB 61 LSSLGIPVNHLEIGSKQVAEIGPQAVGAVKALKALLGALTIVFG 104
XX
XX RESULT 2
XX AAY44458
XX ID AAY44458 standard; protein; 104 AA.
XX AC AAY44458;
XX
XX XX 27-MAR-2000 (first entry)
XX
XX Human lung specific gene protein lng107.
XX
XX Lung Specific Gene; LSG; lng107; human; diagnostic marker; prognosticate;
XX lung cancer; diagnosis.
XX

```

```

XX OS Homo sapiens.
XX PN WO9960160-A1.
XX XX 25-NOV-1999.
XX PD 12-MAY-1999; 99WO-US010344.
XX PF 21-MAY-1998; 98US-0086212P.
XX PR (DIAD-) DIADEXUS LLC.
XX PA
XX PI Yang F, Macina RA, Sun Y;
XX XX WPI; 2000-116320/10.
XX DR N-PSDB; AAZ29723.
XX
XX A new method for diagnosing, monitoring and staging lung cancer.
XX
XX Example 2; Page 38-39; 40pp; English.
XX
XX The present sequence is a lung specific gene (LSG) protein lng107 from
XX human clone ID 586271. The LSG has high level of tissue specificity for
XX lungs and is overexpressed in cancerous tissues. The sequence serves as a
XX diagnostic marker for detecting, monitoring, staging and prognosticating
XX lung cancer. The diagnosis involves comparing levels of LSG in samples
XX obtained from patient and normal control
XX
XX Sequence 104 AA;
XX
XX Query Match 100.0%; Score 502; DB 3; Length 104;
XX Best Local Similarity 100.0%; Pred. No. 1.4e-48;
XX Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MKLAALGLCVALSCHSSAAAFVLSGSAKPVAQPVAALESAAEAGAGTLANPLGTLNPKLL 60
XX DB 1 MKLAALGLCVALSCHSSAAAFVLSGSAKPVAQPVAALESAAEAGAGTLANPLGTLNPKLL 60
XX
XX QY 61 LSSLGIPVNHLEIGSKQVAEIGPQAVGAVKALKALLGALTIVFG 104
XX DB 61 LSSLGIPVNHLEIGSKQVAEIGPQAVGAVKALKALLGALTIVFG 104
XX
XX RESULT 3
XX AAY87288
XX ID AAY87288 standard; protein; 104 AA.
XX AC AAY87288;
XX
XX XX 11-MAY-2000 (first entry)
XX
XX Human signal peptide containing protein HSPP-65 SEQ ID NO:65.
XX
XX Human; signal peptide-containing protein; HSPP; diagnosis; cancer;
XX inflammation; cardiovascular disease; anticancer; anti-inflammatory;
XX antimicrobial; neurotropic; neuroprotective; cardiovascular; hepatotropic;
XX antiasthmatic; gene therapy; cell proliferation; neurological disorder;
XX reproductive disorder; developmental disorder; arteriosclerosis;
XX cirrhosis; psoriasis; acquired immune deficiency syndrome; anaemia;
XX asthma; Crohn's disease; infection; Alzheimer's disease; schizophrenia;
XX Parkinson's disease; Huntington's diseases; ovulatory defect;
XX muscular dystrophy.
XX
XX OS Homo sapiens.
XX PN WO2000000610-A2.
XX XX 06-JAN-2000.
XX PD 25-JUN-1999; 99WO-US014484.
XX PF 26-JUN-1998; 98US-0090762P.
XX PR

```

```

PR 31-JUL-1998; 98US-0094983P.
PR 01-OCT-1998; 98US-0102686P.
PR 11-DEC-1998; 98US-0112129P.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn MR;
PI Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;
PI Bandman O;
XX
XX WPI; 2000-160673/14.
DR N-PSDB; AA298173.
XX
XX New human signal peptide-containing proteins useful in treatment,
PT prevention and diagnosis of e.g. cancer, inflammation and cardiovascular
PT disease.
XX
XX Claim 1; Page 206; 327pp; English.
XX
XX AA298109 to AA298242 encode AA587224 to AA587357 which represent the
CC human signal peptide-containing proteins HSP-1 to HSP-134. HSPs have
CC anticancer, anti-inflammatory, antimicrobial, neurotropic, hepatotropic,
CC neuroprotective, cardiovascular and antidiabetic activities, and can be
CC used in gene therapy. HSPs can be used to treat or prevent disorders
CC associated with decreased activity or function of HSP. Antagonists of
CC HSP are used to treat or prevent disorders associated with increased
CC activity or function of HSP. Such diseases include cell proliferation
CC (including cancer), inflammation, cardiovascular, neurological,
CC reproductive or developmental disorders, (e.g. arteriosclerosis,
CC cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,
CC asthma, Crohn's disease, microbial or other infections, congestive or
CC ischaemic heart disease, Alzheimer's, Parkinson's or Huntington's
CC diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSP
CC nucleic acids can be used for the recombinant production of HSP, for
CC detecting HSP in standard hybridisation and amplification assays (for
CC diagnosis and monitoring), in gene therapy, as antisense, triplex-forming
CC or ribozyme therapeutics, for detecting related sequences or genetic
CC variations, and for chromosomal mapping. HSP are also used to raise
CC specific antibodies (Ab) and to screen for agonists and antagonists
CC (potential therapeutic agents). Ab are used to diagnose, or monitor, HSP
CC -related diseases (in usual immunoassays), as therapeutic antagonists, in
CC competitive drug screens, and for purification of HSP from natural
CC sources
XX
XX Sequence 104 AA;
SQ
Query Match 100.0%; Score 502; DB 3; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.4e-48;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKLAALLGLCVALLSCSSAAAFVGSAPVQAQPVAALESAAEAGAGTLANPLGTLNPLKLL 60
Db 1 MKLAALLGLCVALLSCSSAAAFVGSAPVQAQPVAALESAAEAGAGTLANPLGTLNPLKLL 60
QY 61 LSSLGIPVNHLEGSQKCVAEFGPQAVGAVKALKALLGALTVEG 104
Db 61 LSSLGIPVNHLEGSQKCVAEFGPQAVGAVKALKALLGALTVEG 104
RESULT 4
AAB65280
ID AAB65280 standard; protein: 104 AA.
XX
XX AAB65280;
XX
XX 02-APR-2001 (first entry)
XX
XX Human PRO1245 (UNQ629) protein sequence SEQ ID NO:408.
XX
XX Human; secreted and transmembrane protein; PRO; cytostatic; cell death;
KW cancer; chromosomal mapping; gene mapping; tissue typing;
KW diagnostic assay.
XX

```

```

OS Homo sapiens.
XX
XX WO2000073454-A1.
PN
XX
XX 07-DEC-2000.
PD
XX
XX 30-MAR-2000; 2000WO-US008439.
XX
XX 02-JUN-1999; 99WO-US012252.
XX 23-JUN-1999; 99US-0141037P.
XX 23-JUL-1999; 99US-0143048P.
XX 20-JUL-1999; 99US-0144758P.
XX 26-JUL-1999; 99US-0145688P.
XX 28-JUL-1999; 99US-0146222P.
XX 17-AUG-1999; 99US-0149396P.
XX 15-SEP-1999; 99WO-US021090.
XX 15-SEP-1999; 99WO-US021547.
XX 08-OCT-1999; 99US-0158663P.
XX 30-NOV-1999; 99WO-US028313.
XX 01-DEC-1999; 99WO-US028301.
XX 16-DEC-1999; 99WO-US030095.
XX 20-DEC-1999; 99WO-US030911.
XX 05-JAN-2000; 2000WO-US000219.
XX 06-JAN-2000; 2000WO-US000376.
XX 11-FEB-2000; 2000WO-US003565.
XX 18-FEB-2000; 2000WO-US004341.
XX 22-FEB-2000; 2000WO-US004414.
XX 24-FEB-2000; 2000WO-US004914.
XX 24-FEB-2000; 2000WO-US005004.
XX 02-MAR-2000; 2000WO-US005841.
XX 15-MAR-2000; 2000WO-US006884.
XX 20-MAR-2000; 2000WO-US007377.
XX
XX (GETH ) GENEVTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
XX Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;
XX Grimaldi CJ, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF;
XX Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;
XX Zhang Z;
XX WPI; 2001-032160/04.
XX N-PSDB; AAF44249.
XX
XX PRO polynucleotides used to produce polypeptides used to target bioactive
XX molecules such as toxins, radiolabels or antibodies, to specific cells,
XX to cause targeted cell death.
XX
XX Claim 12; Fig 290; 935pp; English.
XX
XX The present invention describes human secreted and transmembrane PRO
XX proteins. The PRO proteins have cytostatic activity. The PRO proteins can
XX be used for targeted delivery of bioactive molecules, such as toxins,
XX radiolabels or antibodies, that cause cell death. PRO nucleotide
XX sequences, and their fragments, can be used as hybridisation probes, in
XX chromosomal and gene mapping, and in the generation of anti-sense RNA and
XX DNA. They may also be used to produce transgenic animals which are used
XX to develop and screen therapeutically useful reagents. The PRO nucleotide
XX and protein sequence can be used for tissue typing and in treating
XX cancer. Anti-PRO antibodies can be used in diagnostic assays. AAF44270 to
XX AAF44470 represent PCR primers and hybridisation probes used in the
XX isolation of human PRO sequences. AAF44087 to AAF44269 and AAB65154 to
XX AAB65300 represent human PRO polynucleotide and protein sequences given
XX in the exemplification of the present invention
XX
XX Sequence 104 AA;
SQ
Query Match 100.0%; Score 502; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.4e-48;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKLAALLGLCVALLSCSSAAAFVGSAPVQAQPVAALESAAEAGAGTLANPLGTLNPLKLL 60

```

Db 1 MKLALLGLCVALSAAFLVGSAPVAPVVALESAAEAGATLANPLGTLNPLKLL 60  
Qy 61 LSSLGIPVNHLEIGSQKCVAEELGPQAVGAKKALLGALTIVFG 104  
Best Local Similarity 100.0%; Score 502; DB 5; Length 104;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 61 LSSLGIPVNHLEIGSQKCVAEELGPQAVGAKKALLGALTIVFG 104  
RESULT 5  
AAU86141  
ID AAU86141 standard; protein; 104 AA.  
XX  
AC AAU86141;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE Human PRO1245 polypeptide.  
XX  
KW Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;  
KW leukemia; neuronal disorder; stromal disorder; blastocoealic disorder;  
KW inflammatory disorder; immune disorder; angiogenic disorder; cytostatic;  
KW neuroprotective.  
XX  
OS Homo sapiens.  
XX  
PN WO200153486-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 11-FEB-2000; 2000WO-US003565.  
XX  
PR 08-MAR-1999; 99WO-US005028.  
PR 11-MAR-1999; 99US-0123972P.  
PR 11-MAY-1999; 99US-0133459P.  
PR 02-JUN-1999; 99WO-US012252.  
PR 22-JUN-1999; 99US-0140650P.  
PR 22-JUN-1999; 99US-0140653P.  
PR 20-JUL-1999; 99US-0144758P.  
PR 26-JUL-1999; 99US-0145698P.  
PR 28-JUL-1999; 99US-0146222P.  
PR 17-AUG-1999; 99US-0149395P.  
PR 31-AUG-1999; 99US-0151689P.  
PR 01-SEP-1999; 99WO-US020111.  
PR 15-SEP-1999; 99WO-US021090.  
PR 30-NOV-1999; 99WO-US028313.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 05-JAN-2000; 2000WO-US000219.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;  
PI Marsters SA, Pan J, Pitti RM, Roy MA, Smith V, Stone DM;  
PI Watanabe CK, Wood WI;  
XX  
DR WPI; 2002-205567/26.  
DR N-PSDB; ABK40267.  
XX  
PT Thirty five nucleic acids encoding PRO polypeptides, useful for treating  
PT benign or malignant tumors, leukemias and lymphoid malignancies,  
PT inflammatory, angiogenic and immunologic disorders.  
XX  
PS Claim 61; Fig 28; 302pp; English.  
XX  
CC The present invention relates to the isolation of novel human PRO  
CC polypeptides and the polynucleotide sequences encoding them. The PRO  
CC polypeptides, agonists, antagonists or anti-PRO antibodies are useful for  
CC treating benign or malignant tumours (e.g. renal, kidney, bladder,  
CC breast, etc), leukemias and lymphoid malignancies, other disorders such  
CC as neuronal, glial, astrocytal, hypothalamic, glandular, macrophagal,  
CC stromal and blastocoealic disorders, inflammatory, immune and angiogenic  
CC disorders. The polynucleotide sequences are also useful in gene therapy.  
CC AAU86128-AAU86162 represent the human PRO polypeptides of the invention  
XX

SQ Sequence 104 AA;  
Query Match 100.0%; Score 502; DB 5; Length 104;  
Best Local Similarity 100.0%; Pred. No. 1.4e-48;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MKLALLGLCVALSAAFLVGSAPVAPVVALESAAEAGATLANPLGTLNPLKLL 60  
Db 1 MKLALLGLCVALSAAFLVGSAPVAPVVALESAAEAGATLANPLGTLNPLKLL 60  
Qy 61 LSSLGIPVNHLEIGSQKCVAEELGPQAVGAKKALLGALTIVFG 104  
Db 61 LSSLGIPVNHLEIGSQKCVAEELGPQAVGAKKALLGALTIVFG 104  
RESULT 6  
ABU58095  
ID ABU58095 standard; protein; 104 AA.  
XX  
AC ABU58095;  
XX  
DT 14-APR-2003 (first entry)  
XX  
DE Human PRO polypeptide #127.  
XX  
KW Human; PRO; cytostatic; tumour; cancer; breast; lung; stomach; liver;  
KW horse; cow; dog; cat; sheep; pig; goat; rabbit; ADEPT;  
KW antibody-dependent enzyme mediated prodruq therapy.  
XX  
OS Homo sapiens.  
XX  
PN US2003027163-A1.  
XX  
PD 06-FEB-2003.  
XX  
PF 15-NOV-2001; 2001US-00997666.  
XX  
PR 16-JUN-1997; 97US-0049787P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 05-NOV-1997; 97WO-US020069.  
PR 12-NOV-1997; 97US-0065186P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 24-NOV-1997; 97US-0068770P.  
PR 23-FEB-1998; 98US-0075945P.  
PR 20-MAR-1998; 98US-0078910P.  
PR 28-APR-1998; 98US-0083322P.  
PR 07-MAY-1998; 98US-0084600P.  
PR 28-MAY-1998; 98US-0087106P.  
PR 02-JUN-1998; 98US-0087607P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088021P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088026P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088030P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088734P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088742P.  
PR 10-JUN-1998; 98US-0088810P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088858P.  
PR 11-JUN-1998; 98US-0088861P.

PR	11-JUN-1998;	98US-0088876P.	PR	17-AUG-1998;	98US-0096867P.
PR	12-JUN-1998;	98US-0089105P.	PR	17-AUG-1998;	98US-0096891P.
PR	16-JUN-1998;	98US-0089440P.	PR	17-AUG-1998;	98US-0096894P.
PR	16-JUN-1998;	98US-00895112P.	PR	17-AUG-1998;	98US-0096895P.
PR	16-JUN-1998;	98US-00895114P.	PR	17-AUG-1998;	98US-0096897P.
PR	17-JUN-1998;	98US-0089532P.	PR	18-AUG-1998;	98US-0096949P.
PR	17-JUN-1998;	98US-0089538P.	PR	18-AUG-1998;	98US-0096950P.
PR	17-JUN-1998;	98US-0089598P.	PR	18-AUG-1998;	98US-0096959P.
PR	17-JUN-1998;	98US-0089599P.	PR	18-AUG-1998;	98US-0096960P.
PR	17-JUN-1998;	98US-0089600P.	PR	19-AUG-1998;	98US-0097022P.
PR	17-JUN-1998;	98US-0089653P.	PR	19-AUG-1998;	98US-0097111P.
PR	18-JUN-1998;	98US-0089601P.	PR	20-AUG-1998;	98US-0097218P.
PR	18-JUN-1998;	98US-0089907P.	PR	24-AUG-1998;	98US-0097661P.
PR	18-JUN-1998;	98US-0089908P.	PR	26-AUG-1998;	98US-0097952P.
PR	19-JUN-1998;	98US-0089947P.	PR	26-AUG-1998;	98US-0097954P.
PR	19-JUN-1998;	98US-0089948P.	PR	26-AUG-1998;	98US-0097955P.
PR	19-JUN-1998;	98US-0089952P.	PR	26-AUG-1998;	98US-0097971P.
PR	22-JUN-1998;	98US-0090246P.	PR	26-AUG-1998;	98US-0097974P.
PR	22-JUN-1998;	98US-0090252P.	PR	26-AUG-1998;	98US-0097978P.
PR	22-JUN-1998;	98US-0090254P.	PR	26-AUG-1998;	98US-0097979P.
PR	23-JUN-1998;	98US-0090349P.	PR	26-AUG-1998;	98US-0097986P.
PR	23-JUN-1998;	98US-0090355P.	PR	26-AUG-1998;	98US-0098014P.
PR	23-JUN-1998;	98US-0090429P.	PR	31-AUG-1998;	98US-0098535P.
PR	24-JUN-1998;	98US-0090431P.	PR	16-SEP-1998;	98US-0100634P.
PR	24-JUN-1998;	98US-0090435P.	PR	16-SEP-1998;	98US-0100634P.
PR	24-JUN-1998;	98US-0090444P.	PR	17-SEP-1998;	98US-0100858P.
PR	24-JUN-1998;	98US-0090445P.	PR	17-SEP-1998;	98US-0100858P.
PR	24-JUN-1998;	98US-0090472P.	PR	07-OCT-1998;	98US-0100858P.
PR	24-JUN-1998;	98US-0090535P.	PR	01-DEC-1998;	98US-0100858P.
PR	24-JUN-1998;	98US-0090540P.	PR	22-DEC-1998;	98US-0113296P.
PR	24-JUN-1998;	98US-0090542P.	PR	05-JAN-1999;	98US-0113296P.
PR	24-JUN-1998;	98US-0090557P.	PR	08-MAR-1999;	98US-0113296P.
PR	25-JUN-1998;	98US-0090676P.	PR	08-MAR-1999;	98US-0113296P.
PR	25-JUN-1998;	98US-0090678P.	PR	12-MAR-1999;	98US-0113296P.
PR	25-JUN-1998;	98US-0090690P.	PR	02-JUN-1999;	98US-0113296P.
PR	25-JUN-1998;	98US-0090694P.	PR	23-JUN-1999;	98US-0113296P.
PR	25-JUN-1998;	98US-0090695P.	PR	07-JUL-1999;	98US-0143048P.
PR	25-JUN-1998;	98US-0090696P.	PR	20-JUL-1999;	98US-0144758P.
PR	26-JUN-1998;	98US-0090862P.	PR	26-JUL-1999;	98US-0145698P.
PR	26-JUN-1998;	98US-0090863P.	PR	28-JUL-1999;	98US-0145698P.
PR	01-JUL-1998;	98US-0091360P.	PR	17-AUG-1999;	98US-0145698P.
PR	01-JUL-1998;	98US-0091444P.	PR	15-SEP-1999;	98US-0145698P.
PR	02-JUL-1998;	98US-0091478P.	PR	15-SEP-1999;	98US-0145698P.
PR	02-JUL-1998;	98US-0091519P.	PR	08-OCT-1999;	98US-0158663P.
PR	02-JUL-1998;	98US-0091626P.	PR	30-NOV-1999;	98US-0158663P.
PR	02-JUL-1998;	98US-0091628P.	PR	01-DEC-1999;	98US-0158663P.
PR	02-JUL-1998;	98US-0091633P.	PR	01-DEC-1999;	98US-0158663P.
PR	02-JUL-1998;	98US-0091673P.	PR	16-DEC-1999;	98US-0158663P.
PR	02-JUL-1998;	98US-0091978P.	PR	20-DEC-1999;	98US-0158663P.
PR	07-JUL-1998;	98US-0091982P.	PR	05-JAN-2000;	98US-0158663P.
PR	09-JUL-1998;	98US-0092182P.	PR	11-FEB-2000;	98US-0158663P.
PR	10-JUL-1998;	98US-0092472P.	PR	11-FEB-2000;	98US-0158663P.
PR	20-JUL-1998;	98US-0093339P.	PR	22-FEB-2000;	98US-0158663P.
PR	30-JUL-1998;	98US-0094651P.	PR	24-FEB-2000;	98US-0158663P.
PR	04-AUG-1998;	98US-0095282P.	PR	02-MAR-2000;	98US-0158663P.
PR	04-AUG-1998;	98US-0095285P.	PR	10-MAR-2000;	98US-0158663P.
PR	04-AUG-1998;	98US-0095301P.	PR	15-MAR-2000;	98US-0158663P.
PR	04-AUG-1998;	98US-0095302P.	PR	20-MAR-2000;	98US-0158663P.
PR	04-AUG-1998;	98US-0095318P.	PR	30-MAR-2000;	98US-0158663P.
PR	04-AUG-1998;	98US-0095321P.	PR	30-MAR-2000;	98US-0158663P.
PR	04-AUG-1998;	98US-0095325P.	PR	15-MAY-2000;	98US-0158663P.
PR	10-AUG-1998;	98US-0095916P.	PR	22-MAY-2000;	98US-0158663P.
PR	10-AUG-1998;	98US-0095929P.	PR	30-MAY-2000;	98US-0158663P.
PR	10-AUG-1998;	98US-0096012P.	PR	03-JUN-2000;	98US-0158663P.
PR	11-AUG-1998;	98US-0096143P.	PR	23-JUN-2000;	98US-0158663P.
PR	11-AUG-1998;	98US-0096146P.	PR	28-JUL-2000;	98US-0213637P.
PR	12-AUG-1998;	98US-0096329P.	PR	11-AUG-2000;	98US-0213637P.
PR	17-AUG-1998;	98US-0096757P.	PR	23-AUG-2000;	98US-0213637P.
PR	17-AUG-1998;	98US-0096766P.	PR	24-AUG-2000;	98US-0213637P.
PR	17-AUG-1998;	98US-0096768P.	PR	07-SEP-2000;	98US-0213637P.
PR	17-AUG-1998;	98US-0096773P.	PR	07-SEP-2000;	98US-0213637P.
PR	17-AUG-1998;	98US-0096791P.	PR	07-SEP-2000;	98US-0213637P.

Query Match

100.0%

Score 502; DB 6; Length 104;



Novel isolated PRO polypeptides e.g., PRO826, PRO1068, PRO1184, PRO1346 and PRO1375, which stimulate proliferation of stimulated T-lymphocytes are therapeutically useful for enhancing immune response and in cancer treatments.

Claim 12; Fig 290; 648pp; English.

The invention describes an isolated human PRO polypeptide. The PRO polypeptides are useful in detecting PRO polypeptides in a sample, in linking a bioactive molecule to a cell expressing a PRO polypeptide, and in modulating at least one biological activity of a cell expressing a PRO polypeptide. PRO1312 stimulates hypertrophy of neonatal heart and is thus useful for treating cardiac insufficiency disorders. PRO1154 and PRO1186 stimulate adrenal cortical capillary endothelial growth and PRO536, PRO943, PRO828, PRO826, PRO1068 or PRO535, PRO826, PRO819, PRO1126, PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus useful for treating conditions or disorders where angiogenesis would be beneficial, e.g. wound healing and antagonist of this polypeptide are useful for treating cancerous tumours. PRO812 inhibits vascular endothelial growth factor (VEGF) stimulated proliferation of endothelial cells and is thus useful for inhibiting endothelial cell growth in mammals which would be beneficial in inhibiting tumour growth. PRO826, PRO1068, PRO1184, PRO1346 and PRO1375 stimulate proliferation of stimulated T-lymphocytes and are therapeutically useful for enhancing immune response. PRO828, PRO826, PRO1068 or PRO1132 enhance survival of retinal neurons cells (PRO1132 is also enhances survival/proliferation of rod photoreceptor cells) and therefore are useful for treating retinal disorders of injuries, e.g. retinitis pigmentosa, AMD. PRO819, PRO813 and PRO1066 induce proliferation of mammalian kidney mesangial cells, and therefore are useful for treating kidney disorders associated with decreased mesangial cell function such as Berger disease or other nephropathies associated with dermatitis, herpeticiformis or Crohn's disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the proliferation and/or redifferentiation of chondrocytes in culture and are thus useful for treating sports injuries, and arthritis. This is the amino acid sequence of a novel human PRO protein

Sequence 104 AA;

Query Match 100.0%; Score 502; DB 6; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-48;  
 Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKUALLGLCVALLSCSSAAFLVGSAPVQVAALESAAEAGACTLANPLGTLNPKLL 60  
 DB 1 MKUALLGLCVALLSCSSAAFLVGSAPVQVAALESAAEAGACTLANPLGTLNPKLL 60  
 QY 61 LSSLGIPVNHLEIGSKVAEIGPQAVGAVKALKALLGALTVP 104  
 DB 61 LSSLGIPVNHLEIGSKVAEIGPQAVGAVKALKALLGALTVP 104

RESULT 8

ABU82685

ID ABU82685 standard; protein; 104 AA.

XX AC ABU82685,

XX DT 26-JUN-2003 (first entry)

XX DE Human secreted/transmembrane protein PRO1245.

XX KW Human; PRO; secreted protein; transmembrane protein;

XX KW cardiac insufficiency disorders; angiogenesis; wound healing;

XX KW cancerous tumour; immune response; retinal disorder; sight loss;

XX KW retinitis pigmentosa; age-related macular degeneration; AMD;

XX KW kidney disorder; Berger disease; nephropathy; dermatitis; herpeticiformis;

XX KW Crohn's disease; sports injury; arthritis.

XX OS Homo sapiens.

XX XX US2003032023-A1.

XX XX

PD 13-FEB-2003.  
 XX 14-NOV-2001; 2001US-00990711.  
 XX 16-JUN-1997; 97US-0049787P.  
 PR 17-OCT-1997; 97US-0062250P.  
 PR 05-NOV-1997; 97WO-US020069.  
 PR 12-NOV-1997; 97US-0065186P.  
 PR 13-NOV-1997; 97US-0065311P.  
 PR 24-NOV-1997; 97US-0066770P.  
 PR 25-FEB-1998; 98US-0075945P.  
 PR 20-MAR-1998; 98US-0078910P.  
 PR 28-APR-1998; 98US-0083322P.  
 PR 07-MAY-1998; 98US-0084600P.  
 PR 28-MAY-1998; 98US-0087106P.  
 PR 02-JUN-1998; 98US-0087607P.  
 PR 02-JUN-1998; 98US-0087609P.  
 PR 02-JUN-1998; 98US-0087759P.  
 PR 03-JUN-1998; 98US-0087827P.  
 PR 04-JUN-1998; 98US-0088021P.  
 PR 04-JUN-1998; 98US-0088025P.  
 PR 04-JUN-1998; 98US-0088026P.  
 PR 04-JUN-1998; 98US-0088028P.  
 PR 04-JUN-1998; 98US-0088029P.  
 PR 04-JUN-1998; 98US-0088030P.  
 PR 04-JUN-1998; 98US-0088033P.  
 PR 04-JUN-1998; 98US-0088326P.  
 PR 05-JUN-1998; 98US-0088167P.  
 PR 05-JUN-1998; 98US-0088202P.  
 PR 05-JUN-1998; 98US-0088212P.  
 PR 05-JUN-1998; 98US-0088217P.  
 PR 09-JUN-1998; 98US-0088655P.  
 PR 10-JUN-1998; 98US-0088734P.  
 PR 10-JUN-1998; 98US-0088738P.  
 PR 10-JUN-1998; 98US-0088742P.  
 PR 10-JUN-1998; 98US-0088810P.  
 PR 10-JUN-1998; 98US-0088824P.  
 PR 10-JUN-1998; 98US-0088828P.  
 PR 11-JUN-1998; 98US-0088858P.  
 PR 11-JUN-1998; 98US-0088861P.  
 PR 12-JUN-1998; 98US-0088876P.  
 PR 12-JUN-1998; 98US-0089105P.  
 PR 16-JUN-1998; 98US-0089440P.  
 PR 16-JUN-1998; 98US-0089512P.  
 PR 16-JUN-1998; 98US-0089514P.  
 PR 17-JUN-1998; 98US-0089532P.  
 PR 17-JUN-1998; 98US-0089538P.  
 PR 17-JUN-1998; 98US-0089598P.  
 PR 17-JUN-1998; 98US-0089599P.  
 PR 17-JUN-1998; 98US-0089600P.  
 PR 17-JUN-1998; 98US-0089653P.  
 PR 18-JUN-1998; 98US-0089801P.  
 PR 18-JUN-1998; 98US-0089907P.  
 PR 18-JUN-1998; 98US-0089908P.  
 PR 19-JUN-1998; 98US-0089947P.  
 PR 19-JUN-1998; 98US-0089948P.  
 PR 19-JUN-1998; 98US-0089952P.  
 PR 22-JUN-1998; 98US-0090246P.  
 PR 22-JUN-1998; 98US-0090252P.  
 PR 22-JUN-1998; 98US-0090254P.  
 PR 23-JUN-1998; 98US-0090349P.  
 PR 23-JUN-1998; 98US-0090355P.  
 PR 24-JUN-1998; 98US-0090429P.  
 PR 24-JUN-1998; 98US-0090431P.  
 PR 24-JUN-1998; 98US-0090435P.  
 PR 24-JUN-1998; 98US-0090444P.  
 PR 24-JUN-1998; 98US-0090445P.  
 PR 24-JUN-1998; 98US-0090472P.  
 PR 24-JUN-1998; 98US-0090535P.  
 PR 24-JUN-1998; 98US-0090540P.  
 PR 24-JUN-1998; 98US-0090542P.  
 PR 24-JUN-1998; 98US-0090557P.  
 PR 25-JUN-1998; 98US-0090676P.



PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 01-JUL-1998; 98US-0091360P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091519P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091633P.  
PR 02-JUL-1998; 98US-0091646P.  
PR 02-JUL-1998; 98US-0091673P.  
PR 07-JUL-1998; 98US-0091978P.  
PR 07-JUL-1998; 98US-0091982P.  
PR 09-JUL-1998; 98US-0092182P.  
PR 10-JUL-1998; 98US-0092472P.  
PR 20-JUL-1998; 98US-0093339P.  
PR 30-JUL-1998; 98US-0094651P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 04-AUG-1998; 98US-0095285P.  
PR 04-AUG-1998; 98US-0095301P.  
PR 04-AUG-1998; 98US-0095302P.  
PR 04-AUG-1998; 98US-0095318P.  
PR 04-AUG-1998; 98US-0095321P.  
PR 04-AUG-1998; 98US-0095325P.  
PR 10-AUG-1998; 98US-0095916P.  
PR 10-AUG-1998; 98US-0095929P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 11-AUG-1998; 98US-0096143P.  
PR 11-AUG-1998; 98US-0096146P.  
PR 12-AUG-1998; 98US-0096329P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096768P.  
PR 17-AUG-1998; 98US-0096773P.  
PR 17-AUG-1998; 98US-0096791P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096894P.  
PR 17-AUG-1998; 98US-0096895P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096950P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0096960P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 19-AUG-1998; 98US-0097141P.  
PR 20-AUG-1998; 98US-0097218P.  
PR 24-AUG-1998; 98US-0097661P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0097978P.  
PR 26-AUG-1998; 98US-0097979P.  
PR 26-AUG-1998; 98US-0097986P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 31-AUG-1998; 98US-0098525P.  
PR 16-SEP-1998; 98US-0100634P.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-0100858P.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 01-DEC-1998; 98WO-US025108.  
PR 22-DEC-1998; 98US-0113296P.  
PR 05-JAN-1999; 98WO-US000106.  
PR 08-MAR-1999; 98WO-US005028.  
PR 12-MAR-1999; 99US-0123957P.

PR 02-JUN-1999; 99WO-US012252.  
PR 23-JUN-1999; 99US-0141037P.  
PR 07-JUL-1999; 99US-0143048P.  
PR 20-JUL-1999; 99US-0144758P.  
PR 26-JUL-1999; 99US-0145698P.  
PR 28-JUL-1999; 99US-0146222P.  
PR 17-AUG-1999; 99US-0149396P.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 08-OCT-1999; 99US-0158663P.  
PR 30-NOV-1999; 99WO-US028313.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030911.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 15-MAY-2000; 2000WO-US013358.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 23-JUN-2000; 2000US-0213637P.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
  
Query Match 100.0%; Score 502; DB 6; Length 104;  
Best Local Similarity 100.0%; Pred.No. 1.4e-48;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MKLAALLGLCVALSCTSSAAAFVLSAKPVAQPVAALESAAEAGAGTLANPLGTLPKLL 60  
D5 1 MKLAALLGLCVALSCTSSAAAFVLSAKPVAQPVAALESAAEAGAGTLANPLGTLPKLL 60  
  
QY 61 LSSLGIPVNHLEGSQKVAELGPQAVGAVKALKALLGALTTFVG 104  
D5 61 LSSLGIPVNHLEGSQKVAELGPQAVGAVKALKALLGALTTFVG 104  
  
RESULT 9  
AAO19895  
ID AAO19895 standard; protein; 104 AA.  
XX  
AC AAO19895;  
XX  
DT 11-AUG-2003 (first entry)  
XX  
DE Human uteroglobin related protein 2.  
XX  
KW UGSF1; human; mouse; promoter; uteroglobin related protein 1;  
KW respiratory disorder; asthma.  
XX  
OS Homo sapiens.  
XX  
PN WO2003000111-A2.  
XX  
PD 03-JAN-2003.  
XX  
PF 18-JUN-2002; 2002WO-US019456.  
XX  
PR 20-JUN-2001; 2001US-0299828P.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

```
XX Kimura S, Niimi T;
PI WPI; 2003-184004/18.
DR
XX New human UGRPI nucleic acid, useful for diagnosing or predicting a
PT predisposition to develop a respiratory disorder or determining the
PT prognosis of a subject having or suspected of having a respiratory
PT disorder e.g., asthma.
XX
PS Disclosure; Page 79-80; 83pp; English.
XX
CC The present invention provides the human and murine uteroglobin related
CC protein 1 (UGRPI) promoters. The sequences can be used in the diagnosis
CC of and prediction of predisposition to respiratory disorders such as
CC asthma. The present sequence is a protein sequence shown in the
CC exemplification of the invention
XX
SQ Sequence 104 AA;
Query Match 100.0%; Score 502; DB 6; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.4e-48; Mismatches 0; Gaps 0;
Matches 104; Conservative 0; Indels 0;
Oy 1 MKLAALGLCVLSCSSAAFLVGSAPVAQPVAALESAAEAGTLANPLGTLNPLKL 60
Db 1 MKLAALGLCVLSCSSAAFLVGSAPVAQPVAALESAAEAGTLANPLGTLNPLKL 60
Oy 61 LSSLGIPVNHLEGSQKCVNELGPQAVGAVKALKALGALTVEG 104
Db 61 LSSLGIPVNHLEGSQKCVNELGPQAVGAVKALKALGALTVEG 104
RESULT 10
ABU0604
ID ABU0604 standard; protein; 104 AA.
XX
AC ABU0604;
XX
DT 01-MAY-2003 (first entry)
XX
DE Human secreted/transmembrane protein, #163.
XX
KW Human; PRO; secreted; transmembrane; signal peptide; pharmaceutical;
KW diagnostic; therapeutic; gene therapy.
XX
OS Homo sapiens.
XX
PN US2002160384-A1.
XX
FD 31-OCT-2002.
XX
FF 14-NOV-2001; 2001US-00992598.
XX
PR 16-JUN-1997; 97US-0049787P.
PR 17-OCT-1997; 97US-0062250P.
PR 05-NOV-1997; 97MO-US020069.
PR 12-NOV-1997; 97US-0055186P.
PR 13-NOV-1997; 97US-0085311P.
PR 24-NOV-1997; 97US-0066770P.
PR 25-FEB-1998; 98US-0075945P.
PR 20-MAR-1998; 98US-0078910P.
PR 28-APR-1998; 98US-0083322P.
PR 07-MAY-1998; 98US-0084600P.
PR 28-MAY-1998; 98US-0087106P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087659P.
PR 03-JUN-1998; 98US-0087759P.
PR 04-JUN-1998; 98US-0088021P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088026P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088030P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088036P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088734P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088742P.
PR 10-JUN-1998; 98US-0088810P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088856P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089105P.
PR 12-JUN-1998; 98US-0089440P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089532P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089599P.
PR 17-JUN-1998; 98US-0089600P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089801P.
PR 18-JUN-1998; 98US-0089907P.
PR 18-JUN-1998; 98US-0089908P.
PR 18-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 02-JUN-1999; 99WO-US012252.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 16-DEC-1999; 99WO-US030911.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US005841.
PR 15-MAR-2000; 2000WO-US005884.
PR 20-MAR-2000; 2000WO-US007377.
PR 30-MAR-2000; 2000WO-US008439.
PR 15-MAY-2000; 2000WO-US013358.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-JUN-2001; 2001WO-US015800.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUN-2001; 2001WO-US021735.
```

```
PR 28-AUG-2001; 2001US-00941992.
XX (GETH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Botstein D, Deanovers L, Eaton DL,
PI Perrata N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ,
PI Grimaldi JC, Gruney AL, Kljavin LV, Napier MA, Pan J, Paoni NF,
PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI,
PI Zhang Z;
XX
XX WPI; 2003-288106/28.
DR N-PSDB; ABX90363.
XX
XX New transmembrane polypeptides and nucleic acids encoding the
PI polypeptides, useful in gene therapy, in chromosome identification, as
PI chromosome markers, or in generating probes.
XX
XX Claim 12; Fig 290; 650pp; English.
XX
XX The invention discloses isolated PRO secreted/transmembrane polypeptides
CC comprising a sequence without signal peptide and the nucleic acid
CC encoding them. The polypeptides can be used to raise antibodies that
CC specifically bind to the PRO polypeptide, for linking a bioactive
CC molecule to a cell expressing a PRO protein and for modulating at least
CC one biological activity of a cell. The PRO polypeptides or
CC polynucleotides are also useful in gene therapy, in chromosome
CC identification, as chromosome markers, or in generating probes. The PRO
CC polypeptides are useful as molecular markers for protein electrophoresis,
CC and the isolated nucleic acids may be used for recombinantly expressing
CC these markers. The PRO polypeptides and nucleic acids may also be used in
CC tissue typing. Anti-PRO antibodies are useful in diagnostic assays for
CC PRO, and in affinity purification of PRO from recombinant cell culture or
CC natural sources. The sequences presented in ABU60478-ABU60624 are the PRO
CC polynucleotides of the invention. Note: The sequence data for this patent
CC is also available in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX Sequence 104 AA;
SQ
Query Match 100.0%; Score 502; DB 6; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.4e-48;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKLAALLGLCVALLSCSSAAAFVLSGAKPVAQPVAALESAAEAGAGTLANPLGTLNPKLL 60
DB 1 MKLAALLGLCVALLSCSSAAAFVLSGAKPVAQPVAALESAAEAGAGTLANPLGTLNPKLL 60
QY 61 LSSGLIPVNHLEIGSKCVAEELGPQAVGAKVALLGALTIVFG 104
DB 61 LSSGLIPVNHLEIGSKCVAEELGPQAVGAKVALLGALTIVFG 104
RESULT 11
ABU13986
ID ABU13986 standard; protein; 104 AA.
XX
XX AC ABU13986;
XX
XX DT 26-FEB-2003 (first entry)
XX
XX DE Human PRO1245 polypeptide.
XX
XX KW Human; PRO polypeptide; secreted protein; transmembrane protein;
KW genetic disorder; antibacterial; immunosuppressive.
XX
XX OS Homo sapiens.
XX
XX PN US2002103125-A1.
XX
XX PD 01-AUG-2002.
XX
XX PF 20-NOV-2001; 2001US-00989731.
XX
```

16-JUN-1997; 97US-0049787P.  
17-OCT-1997; 97US-0062250P.  
05-NOV-1997; 97WO-US020069.  
12-NOV-1997; 97US-0065186P.  
13-NOV-1997; 97US-0065311P.  
24-NOV-1997; 97US-0066770P.  
25-FEB-1998; 98US-0075945P.  
20-MAR-1998; 98US-0078910P.  
28-APR-1998; 98US-0083322P.  
07-MAY-1998; 98US-0084600P.  
28-MAY-1998; 98US-0087106P.  
02-JUN-1998; 98US-0087607P.  
02-JUN-1998; 98US-0087609P.  
02-JUN-1998; 98US-0087759P.  
03-JUN-1998; 98US-0087827P.  
04-JUN-1998; 98US-0088021P.  
04-JUN-1998; 98US-0088025P.  
04-JUN-1998; 98US-0088026P.  
04-JUN-1998; 98US-0088028P.  
04-JUN-1998; 98US-0088029P.  
04-JUN-1998; 98US-0088030P.  
04-JUN-1998; 98US-0088033P.  
04-JUN-1998; 98US-0088326P.  
05-JUN-1998; 98US-0088167P.  
05-JUN-1998; 98US-0088202P.  
05-JUN-1998; 98US-0088212P.  
05-JUN-1998; 98US-0088217P.  
09-JUN-1998; 98US-0088655P.  
10-JUN-1998; 98US-0088734P.  
10-JUN-1998; 98US-0088738P.  
10-JUN-1998; 98US-0088742P.  
10-JUN-1998; 98US-0088810P.  
10-JUN-1998; 98US-0088824P.  
10-JUN-1998; 98US-0088826P.  
11-JUN-1998; 98US-0088858P.  
11-JUN-1998; 98US-0088861P.  
11-JUN-1998; 98US-0088876P.  
12-JUN-1998; 98US-0089105P.  
16-JUN-1998; 98US-0089440P.  
16-JUN-1998; 98US-0089512P.  
16-JUN-1998; 98US-0089514P.  
17-JUN-1998; 98US-0089532P.  
17-JUN-1998; 98US-0089538P.  
17-JUN-1998; 98US-0089598P.  
17-JUN-1998; 98US-0089599P.  
17-JUN-1998; 98US-0089600P.  
17-JUN-1998; 98US-0089653P.  
18-JUN-1998; 98US-0089801P.  
18-JUN-1998; 98US-0089907P.  
18-JUN-1998; 98US-0089908P.  
16-SEP-1998; 98WO-US019330.  
17-SEP-1998; 98WO-US019437.  
07-OCT-1998; 98WO-US021141.  
01-DEC-1998; 98WO-US025108.  
05-JAN-1999; 99WO-US000106.  
08-MAR-1999; 99WO-US005028.  
02-JUN-1999; 99WO-US012252.  
15-SEP-1999; 99WO-US021090.  
15-SEP-1999; 99WO-US021547.  
30-NOV-1999; 99WO-US028313.  
01-DEC-1999; 99WO-US028301.  
01-DEC-1999; 99WO-US028634.  
16-DEC-1999; 99WO-US030095.  
20-DEC-1999; 99WO-US030911.  
06-JAN-2000; 2000WO-US000219.  
06-JAN-2000; 2000WO-US000376.  
11-FEB-2000; 2000WO-US003565.  
18-FEB-2000; 2000WO-US004341.  
22-FEB-2000; 2000WO-US004414.  
24-FEB-2000; 2000WO-US004914.  
24-FEB-2000; 2000WO-US005004.  
02-MAR-2000; 2000WO-US005841.  
10-MAR-2000; 2000WO-US006319.

PR 15-MAR-2000; 2000WO-US006884.  
 PR 20-MAR-2000; 2000WO-US007377.  
 PR 30-MAR-2000; 2000WO-US008439.  
 PR 15-MAY-2000; 2000WO-US013358.  
 PR 17-MAY-2000; 2000WO-US013705.  
 PR 22-MAY-2000; 2000WO-US014042.  
 PR 30-MAY-2000; 2000WO-US014941.  
 PR 02-JUN-2000; 2000WO-US015264.  
 PR 28-JUL-2000; 2000WO-US020710.  
 PR 11-AUG-2000; 2000WO-US022031.  
 PR 23-AUG-2000; 2000WO-US023522.  
 PR 24-AUG-2000; 2000WO-US023328.  
 PR 08-NOV-2000; 2000WO-US030952.  
 PR 01-DEC-2000; 2000WO-US032678.  
 PR 28-FEB-2001; 2001WO-US006520.  
 PR 01-JUN-2001; 2001WO-US017800.  
 PR 20-JUN-2001; 2001WO-US019692.  
 PR 29-JUN-2001; 2001WO-US021066.  
 PR 09-JUL-2001; 2001WO-US021735.  
 PR 28-AUG-2001; 2001US-00941992.  
 (GETH ) GENENTECH LTD.

WI: 2003-102117/09.  
 N-PSDB; ABX64209.

Novel secreted and transmembrane polypeptide for modulating biological activity of cell expressing the polypeptide, identifying agonists or antagonists of polypeptide, and as molecular weight markers.

Claim 12; Fig 290; 64pp; English.

The present invention relates to the isolation of novel human PRO polypeptides, and the polynucleotide sequences encoding them. The PRO polypeptides are secreted and transmembrane proteins. The PRO polypeptides are useful for detecting other PRO polypeptides, for linking bioactive molecules to cells expressing PRO polypeptides, for modulating biological activities of cells expressing PRO polypeptides, and for for identifying agonists or antagonists. The polynucleotide sequences encoding PRO polypeptides are useful as hybridisation probes in chromosome and gene mapping, in the generation of antisense RNA and DNA, in the preparation of PRO polypeptides, for generating transgenic animals or knockout animals, to construct hybridisation probes for mapping the gene which encodes the PRO polypeptide, and for the genetic analysis of individuals with genetic disorders, in gene therapy, for chromosome identification, as chromosome markers, and for generating probes for PCR, Northern analysis, Southern analysis and Western analysis. ABU13860-ABU14006 represent the human PRO polypeptides of the invention. Note: The sequence data for this patent was obtained in electronic format directly from the USPTO web site at [seqdata.uspto.gov/psipdbEntry.html](http://seqdata.uspto.gov/psipdbEntry.html)

Sequence 104 AA;

Query Match 100.0%; Score 502; DB 6; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-48;  
 Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLAALGLCVALSAAAFLVGSAAKPVAAQPVAALESAAEAGAGTLANPLGLNPLKLL 60  
 DB 1 MKLAALGLCVALSAAAFLVGSAAKPVAAQPVAALESAAEAGAGTLANPLGLNPLKLL 60  
 QY 61 LSSLGIPVNHLEGSQKCVAEFGQAVGAVKALKALLGALTTFG 104  
 DB 61 LSSLGIPVNHLEGSQKCVAEFGQAVGAVKALKALLGALTTFG 104

RESULT 12

ABU72571  
 ID ABU72571 standard; protein; 104 AA.  
 XX AC ABU72571;  
 XX DT 17-JUN-2003 (first entry)  
 XX DE Novel human secreted and transmembrane protein PRO1245.  
 XX KW Human; secreted and transmembrane protein; cytostatic; anti-HIV; virucide; hepatotropic; antiinflammatory; neuroprotective; gene therapy; PRO; pharmaceutical; diagnostic; biosensor; bioreactor; malignancy; cancer; ovarian cancer; colorectal cancer; Kaposi's sarcoma; leukaemia; lymphoma; hepatitis B; multiple sclerosis; Crohn's disease; drug screening.  
 XX OS Homo sapiens.  
 XX FN US2003003531-A1.  
 XX PD 02-JAN-2003.  
 XX PF 19-NOV-2001; 2001US-00989734.  
 XX PR 16-JUN-1997; 97US-0049787P.  
 PR 17-OCT-1997; 97US-0062250P.  
 PR 05-NOV-1997; 97WO-US020089.  
 PR 12-NOV-1997; 97US-0065186P.  
 PR 13-NOV-1997; 97US-0065311P.  
 PR 24-NOV-1997; 97US-0066770P.  
 PR 25-FEB-1998; 98US-0075945P.  
 PR 20-MAR-1998; 98US-0078910P.  
 PR 28-APR-1998; 98US-0083322P.  
 PR 07-MAY-1998; 98US-0084600P.  
 PR 28-MAY-1998; 98US-0087106P.  
 PR 02-JUN-1998; 98US-0087607P.  
 PR 02-JUN-1998; 98US-0087609P.  
 PR 03-JUN-1998; 98US-0087759P.  
 PR 04-JUN-1998; 98US-0087827P.  
 PR 04-JUN-1998; 98US-0088021P.  
 PR 04-JUN-1998; 98US-0088025P.  
 PR 04-JUN-1998; 98US-0088026P.  
 PR 04-JUN-1998; 98US-0088028P.  
 PR 04-JUN-1998; 98US-0088029P.  
 PR 04-JUN-1998; 98US-0088030P.  
 PR 04-JUN-1998; 98US-0088033P.  
 PR 04-JUN-1998; 98US-0088326P.  
 PR 05-JUN-1998; 98US-0088167P.  
 PR 05-JUN-1998; 98US-0088202P.  
 PR 05-JUN-1998; 98US-0088212P.  
 PR 05-JUN-1998; 98US-0088217P.  
 PR 09-JUN-1998; 98US-0088655P.  
 PR 10-JUN-1998; 98US-0088734P.  
 PR 10-JUN-1998; 98US-0088738P.  
 PR 10-JUN-1998; 98US-0088742P.  
 PR 10-JUN-1998; 98US-0088810P.  
 PR 10-JUN-1998; 98US-0088824P.  
 PR 10-JUN-1998; 98US-0088826P.  
 PR 11-JUN-1998; 98US-0088858P.  
 PR 11-JUN-1998; 98US-0088961P.  
 PR 11-JUN-1998; 98US-0088976P.  
 PR 12-JUN-1998; 98US-0089105P.  
 PR 16-JUN-1998; 98US-0089440P.  
 PR 16-JUN-1998; 98US-0089512P.  
 PR 16-JUN-1998; 98US-0089514P.  
 PR 17-JUN-1998; 98US-0089532P.  
 PR 17-JUN-1998; 98US-0089538P.  
 PR 17-JUN-1998; 98US-0089598P.  
 PR 17-JUN-1998; 98US-0089599P.  
 PR 17-JUN-1998; 98US-0089600P.  
 PR 17-JUN-1998; 98US-0089653P.  
 PR 18-JUN-1998; 98US-0089801P.  
 PR 18-JUN-1998; 98US-0089907P.

PR 18-JUN-1998; 98US-0089908P.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 99WO-US000106.  
PR 08-MAR-1999; 99WO-US005028.  
PR 02-JUN-1999; 99WO-US012252.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 30-NOV-1999; 99WO-US028313.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030911.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 15-MAY-2000; 2000WO-US013358.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 02-JUN-2000; 2000WO-US014941.  
PR 30-JUN-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 28-AUG-2001; 2001WO-US0241992.  
PR 28-AUG-2001; 2001US-00997428.  
PR 15-NOV-2001; 2001US-00997428.  
PR 16-JUN-1997; 97US-0049787P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 05-NOV-1997; 97WO-US020069.  
PR 12-NOV-1997; 97US-0065186P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 24-NOV-1997; 97US-0066770P.  
PR 25-FEB-1998; 98US-0075945P.  
PR 20-MAR-1998; 98US-0078310P.  
PR 28-APR-1998; 98US-0083322P.  
PR 07-MAY-1998; 98US-0084600P.  
PR 28-MAY-1998; 98US-0087106P.  
PR 02-JUN-1998; 98US-0087607P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088021P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088026P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088030P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.

(GETH ) GENENTECH INC.

PR Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL,  
PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ,  
PI Grimaldi JC, Gurney AL, Kiljavin IJ, Napier MA, Pan J, Paoni NF,  
PI Roy MA, Stewart TA, Tamas D, Watanabe CK, Williams PM, Wood WI,  
PI Zhang Z;

DR WPI; 2003-352829/33.  
DR N-PSDB; ACA64431.

PR New genes and secreted and transmembrane polypeptides (e.g. PRO183 or  
PT PRO184), useful for treating or diagnosing e.g. ovarian cancer, Kaposi's  
PT sarcoma, leukemia, lymphoma, hepatitis B, multiple sclerosis or Crohn's  
PT disease.

XX Claim 12; Fig 290; 663pp; English.

XX The invention describes a new isolated nucleic acid molecule comprising  
CC the full length coding sequence of the DNA deposited with the American  
CC Type Culture Collection (e.g. ATCC Deposit No. 209621, 552-PTA, 819-PTA,  
CC 209439, 203135, etc.) or a sequence with at least 80% identity to a DNA  
CC encoding a PRO polypeptide. The PRO polypeptides or polynucleotides are  
CC useful as pharmaceuticals, diagnostics, biosensors or bioreactors. These  
CC are particularly useful for detecting or treating e.g. malignancies or  
CC cancers (e.g. ovarian cancer, colorectal cancer, Kaposi's sarcoma,  
CC leukemia or lymphoma), hepatitis B, multiple sclerosis, or Crohn's  
CC disease in mammals. The PRO polypeptides are useful in drug screening,

CC particularly as targets for therapeutic intervention in these diseases,  
CC and in the diagnostic determination of the presence of these diseases.  
CC The PRO polypeptides are also useful as molecular weight markers, or for  
CC chromosome identification. The PRO genes are useful as hybridisation  
CC probes, or for screening libraries of human cDNA, genomic DNA or mRNA.  
CC The PRO genes may also be used in gene therapy, particularly for a novel  
CC replacing a defective gene. This is the amino acid sequence of a novel  
CC human secreted and transmembrane PRO polypeptide

XX Sequence 104 AA;

Query Match 100.0%; Score 502; DB 6; Length 104;  
Best Local Similarity 100.0%; Pred. No. 1.4e-48;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MKLAALLGLCVALLSCSSAAAFVLSAKPVAQPVAALESAAEAGAGTLANPLKLL 60  
Db 1 MKLAALLGLCVALLSCSSAAAFVLSAKPVAQPVAALESAAEAGAGTLANPLKLL 60

Oy 61 LSSIGIPVNHLEGSOKVAVELGPOAVGAVKALKALLGALTTFVG 104  
Db 61 LSSIGIPVNHLEGSOKVAVELGPOAVGAVKALKALLGALTTFVG 104

RESULT 13

ABUS9320  
ID ABUS9320 standard; protein; 104 AA.

XX AC ABUS9320;

XX DT 22-APR-2003 (first entry)

XX DE Human secreted/transmembrane protein, #163.

XX KW Human; PRO; secreted; transmembrane; pharmaceutical; diagnostic;  
KW biosensor; bioreactor; tumour; therapeutic; gene therapy;  
KW tumour-associated antigenic target; TAT; ADEPT;  
KW antibody-dependent enzyme mediated prodrug therapy; cytostatic.

OS Homo sapiens.

XX FN US2003027162-A1.

XX PD 06-FEB-2003.

XX PF 15-NOV-2001; 2001US-00997428.

PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088734P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088742P.  
PR 10-JUN-1998; 98US-0088810P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 11-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088858P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089440P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089532P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089593P.  
PR 17-JUN-1998; 98US-0089600P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089801P.  
PR 18-JUN-1998; 98US-0089907P.  
PR 18-JUN-1998; 98US-0089982P.  
PR 19-JUN-1998; 98US-0089947P.  
PR 19-JUN-1998; 98US-0089948P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 23-JUN-1998; 98US-0090349P.  
PR 23-JUN-1998; 98US-0090355P.  
PR 24-JUN-1998; 98US-0090423P.  
PR 24-JUN-1998; 98US-0090431P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090445P.  
PR 24-JUN-1998; 98US-0090472P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 24-JUN-1998; 98US-0090542P.  
PR 24-JUN-1998; 98US-0090557P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-0090862P.  
PR 01-JUL-1998; 98US-0090863P.  
PR 01-JUL-1998; 98US-0091360P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091519P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091633P.  
PR 02-JUL-1998; 98US-0091646P.  
PR 02-JUL-1998; 98US-0091673P.  
PR 07-JUL-1998; 98US-0091978P.  
PR 07-JUL-1998; 98US-0091982P.  
PR 09-JUL-1998; 98US-0092182P.  
PR 10-JUL-1998; 98US-0092472P.  
PR 30-JUL-1998; 98US-0093339P.  
PR 04-AUG-1998; 98US-0094651P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 04-AUG-1998; 98US-0095285P.  
PR 04-AUG-1998; 98US-0095301P.  
PR 04-AUG-1998; 98US-0095302P.  
PR 04-AUG-1998; 98US-0095318P.  
PR 04-AUG-1998; 98US-0095321P.  
PR 04-AUG-1998; 98US-0095325P.  
PR 10-AUG-1998; 98US-0095916P.

PR 10-AUG-1998; 98US-0095929P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 11-AUG-1998; 98US-0096143P.  
PR 11-AUG-1998; 98US-0096146P.  
PR 12-AUG-1998; 98US-0096329P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096773P.  
PR 17-AUG-1998; 98US-0096791P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096894P.  
PR 17-AUG-1998; 98US-0096895P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096950P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0096960P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 19-AUG-1998; 98US-0097141P.  
PR 20-AUG-1998; 98US-0097218P.  
PR 24-AUG-1998; 98US-0097661P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0097978P.  
PR 26-AUG-1998; 98US-0097979P.  
PR 26-AUG-1998; 98US-0097986P.  
PR 31-AUG-1998; 98US-0098014P.  
PR 31-AUG-1998; 98US-0098525P.  
PR 16-SEP-1998; 98US-0100634P.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98US-0100858P.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 01-DEC-1998; 98WO-US025108.  
PR 22-DEC-1998; 98US-0113296P.  
PR 05-JAN-1999; 98WO-US000106.  
PR 08-MAR-1999; 98WO-US005028.  
PR 12-MAR-1999; 99US-0123957P.  
PR 02-JUN-1999; 99WO-US012252.  
PR 23-JUL-1999; 99US-0141037P.  
PR 07-JUL-1999; 99US-0143048P.  
PR 20-JUL-1999; 99US-0144758P.  
PR 26-JUL-1998; 99US-0145698P.  
PR 28-JUL-1999; 99US-0146222P.  
PR 17-AUG-1999; 99US-0149396P.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 08-OCT-1999; 99US-0158663P.  
PR 30-NOV-1999; 99WO-US028313.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028324.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030911.  
PR 05-JAN-2000; 2000WO-US0000219.  
PR 06-JAN-2000; 2000WO-US0000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 15-MAY-2000; 2000WO-US013358.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.

```
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 23-JUN-2000; 2000US-0213637P.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.

Query Match 100.0%; Score 502; DB 6; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.4e-48;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKLAALIGLCVAVSCSAAAFVLSGAKPVAOPVAALSAEAGAGTLANPLGTLNPKLL 50
Db 1 MKLAALIGLCVAVSCSAAAFVLSGAKPVAOPVAALSAEAGAGTLANPLGTLNPKLL 50

Qy 61 LSSLGIPVNHIEGSKVCAELGQVAVGAVKALKALIGALTVPF 104
Db 61 LSSLGIPVNHIEGSKVCAELGQVAVGAVKALKALIGALTVPF 104

RESULT 14
ABO26017
ID ABO26017 standard; protein; 104 AA.
AC ABO26017;
DT 10-SEP-2003 (first entry)
DE Human PRO1245 polypeptide.
XX
XX Human; PRO polypeptide; secreted protein; transmembrane protein;
KW genetic disorder; antibacterial; immunosuppressive.
XX
OS Homo sapiens.
XX
XX US2002127576-A1.
XX 12-SEP-2002.
XX
XX 14-NOV-2001; 2001US-00991073.
XX
XX 16-JUN-1997; 97US-0049787P.
XX 17-OCT-1997; 97US-0062250P.
XX 05-NOV-1997; 97WO-US020069.
XX 12-NOV-1997; 97US-0065186P.
XX 13-NOV-1997; 97US-0065311P.
XX 24-NOV-1997; 97US-0066770P.
XX 25-FEB-1998; 98US-0075945P.
XX 20-MAR-1998; 98US-0076910P.
XX 28-APR-1998; 98US-0083322P.
XX 07-MAY-1998; 98US-0084600P.
XX 28-MAY-1998; 98US-0087106P.
XX 02-JUN-1998; 98US-0087607P.
XX 02-JUN-1998; 98US-0087609P.
XX 02-JUN-1998; 98US-0087759P.
XX 03-JUN-1998; 98US-0087827P.
XX 04-JUN-1998; 98US-0088021P.
XX 04-JUN-1998; 98US-0088025P.
XX 04-JUN-1998; 98US-0088026P.
XX 04-JUN-1998; 98US-0088028P.
XX 04-JUN-1998; 98US-0088030P.
XX 04-JUN-1998; 98US-0088033P.
XX 04-JUN-1998; 98US-0088326P.
XX 05-JUN-1998; 98US-0088167P.
XX 05-JUN-1998; 98US-0088202P.
XX 05-JUN-1998; 98US-0088212P.
XX 05-JUN-1998; 98US-0088217P.
XX 09-JUN-1998; 98US-0088555P.
XX 10-JUN-1998; 98US-0088734P.
XX 10-JUN-1998; 98US-0088738P.
XX 10-JUN-1998; 98US-0088742P.
```

```
PR 10-JUN-1998; 98US-0088810P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088858P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089105P.
PR 12-JUN-1998; 98US-0089440P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089532P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089599P.
PR 17-JUN-1998; 98US-0089600P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089801P.
PR 18-JUN-1998; 98US-0089907P.
PR 18-JUN-1998; 98US-0089908P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019437.
PR 07-OCT-1998; 98WO-US021141.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 02-JUN-1999; 99WO-US012252.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 06-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 22-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 10-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 30-MAR-2000; 2000WO-US008439.
PR 15-MAY-2000; 2000WO-US013358.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 11-AUG-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-JUN-2001; 2001WO-US017800.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUL-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 28-AUG-2001; 2001US-00941992.
XX
XX (GETH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Botstein D, Deenoyers L, Eaton DL,
PI Ferrarini N, Fong S, Gerber H, Gertsen ME, Goddard A, Godowski PJ,
PI Grimaldi JC, Gurney AL, Kijavini IU, Napier MA, Pan J, Paoni NF,
PI Roy MA, Stewart TA, Tamas D, Watanabe CK, Williams PM, Wood WI,
XX Zhang Z;
XX WPI; 2003-340824/32.
DR N-PDSB; ACD44399.
```

XX Novel isolated PRO polypeptides e.g., PRO826, PRO1068, PRO1184, PRO1346  
PT and PRO1375, which stimulate proliferation of stimulated T-lymphocytes  
PT and are therapeutically useful for enhancing immune responses.  
XX Claim 12; Fig 290; 661pp; English.  
XX The present invention relates to the isolation of novel human PRO  
CC polypeptides, and the polynucleotide sequences encoding them. The PRO  
CC polypeptides are secreted and transmembrane proteins. The PRO  
CC polypeptides are useful for detecting other PRO polypeptides, for linking  
CC bioactive molecules to cells expressing PRO polypeptides, for modulating  
CC biological activities of cells expressing PRO polypeptides, and for for  
CC identifying agonists or antagonists. The polynucleotide sequences  
CC encoding PRO polypeptides are useful as hybridisation probes, in  
CC chromosome and gene mapping, in the generation of antisense RNA and DNA,  
CC in the preparation of PRO polypeptides, for generating transgenic animals  
CC or knockout animals, to construct hybridisation probes for mapping the  
CC gene which encodes the PRO polypeptide, and for the genetic analysis of  
CC individuals with genetic disorders, in gene therapy, for chromosome  
CC identification, as chromosome markers, and for generating probes for PCR,  
CC Northern analysis, Southern analysis and Western analysis. ABO2891-  
CC ABO26037 represent the human PRO polypeptides of the invention. Note: The  
CC sequence data for this patent was obtained in electronic format directly  
CC from the USPTO web site at seqdata.uspto.gov/psipdbEntry.html  
XX Sequence 104 AA;  
SQ

Query Match 100.0%; Score 502; D8 6; Length 104;  
Best Local Similarity 100.0%; Pred. NO. 1.4e-48;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLALLGLCVALSCTSSAAFLVGSAPVAPVVALESAAEAGAGTIANPLGLTLPKLL 60  
Db 1 MKLALLGLCVALSCTSSAAFLVGSAPVAPVVALESAAEAGAGTIANPLGLTLPKLL 60

QY 61 LSSLGIPVNHLEIGSQKVAELGPQAVKALKALIGALTVP 104  
Db 61 LSSLGIPVNHLEIGSQKVAELGPQAVKALKALIGALTVP 104

RESULT 15  
ABUS9026  
ID ABUS9026 standard; protein; 104 AA.  
XX AC ABUS9026;  
XX DT 16-APR-2003 (first entry)  
XX DE Human secreted/transmembrane protein, #163.  
XX KW Human; PRO; secreted; transmembrane; signal peptide; pharmaceutical;  
KW diagnostic; biosensor; bioreactor; tumour; therapeutic; colon cancer;  
KW lung cancer; breast cancer; cancer; gene therapy.  
XX OS Homo sapiens.  
XX PN US2002142961-A1.  
XX PD 03-OCT-2002.  
XX PF 19-NOV-2001; 2001US-00989721.  
XX PR 16-JUN-1997; 97US-0049787P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 05-NOV-1997; 97WO-US020069.  
PR 12-NOV-1997; 97US-0065186P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 24-NOV-1997; 97US-0066770P.  
PR 25-FEB-1998; 98US-0075945P.  
PR 20-MAR-1998; 98US-0078910P.  
PR 28-APR-1998; 98US-0083322P.  
PR 07-MAY-1998; 98US-0084600P.

PR 28-MAY-1998; 98US-0087106P.  
PR 02-JUN-1998; 98US-0087607P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087753P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088021P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088026P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088036P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088734P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088742P.  
PR 10-JUN-1998; 98US-0088810P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088858P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089440P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089554P.  
PR 17-JUN-1998; 98US-0089532P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089599P.  
PR 17-JUN-1998; 98US-0089600P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089801P.  
PR 18-JUN-1998; 98US-0089907P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 98WO-US000106.  
PR 08-MAR-1999; 98WO-US005028.  
PR 02-JUN-1999; 99WO-US012252.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 30-NOV-1999; 99WO-US028313.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030511.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 15-MAY-2000; 2000WO-US013358.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.



radiolabel, antibody; cell death; tissue typing; gene therapy;  
cytostatic; chromosome mapping; gene mapping; transgenic animal;  
knockout animal; immunohistochemical staining.

KW	radiolabel, antibody; cell death; tissue typing; gene therapy;
KW	cytostatic; chromosome mapping; gene mapping; transgenic animal;
KW	knockout animal; immunohistochemical staining.
XX	Homo sapiens.
OS	US2003022187-A1.
XX	30-JAN-2003.
XX	14-NOV-2001; 2001US-00993567.
XX	16-JUN-1997; 97US-0049787P.
XX	17-OCT-1997; 97US-0062250P.
XX	05-NOV-1997; 97WO-US020069.
XX	13-NOV-1997; 97US-0065186P.
XX	13-NOV-1997; 97US-0065311P.
XX	24-NOV-1997; 97US-0066770P.
XX	25-FEB-1998; 98US-0075945P.
XX	20-MAR-1998; 98US-0078910P.
XX	28-APR-1998; 98US-0083322P.
XX	07-MAY-1998; 98US-0084600P.
XX	28-MAY-1998; 98US-0087106P.
XX	02-JUN-1998; 98US-0087607P.
XX	02-JUN-1998; 98US-0087609P.
XX	02-JUN-1998; 98US-0087759P.
XX	03-JUN-1998; 98US-0087827P.
XX	04-JUN-1998; 98US-0088021P.
XX	04-JUN-1998; 98US-0088025P.
XX	04-JUN-1998; 98US-0088026P.
XX	04-JUN-1998; 98US-0088028P.
XX	04-JUN-1998; 98US-0088029P.
XX	04-JUN-1998; 98US-0088030P.
XX	04-JUN-1998; 98US-0088033P.
XX	04-JUN-1998; 98US-0088326P.
XX	05-JUN-1998; 98US-0088167P.
XX	05-JUN-1998; 98US-0088202P.
XX	05-JUN-1998; 98US-0088212P.
XX	05-JUN-1998; 98US-0088217P.
XX	09-JUN-1998; 98US-0088655P.
XX	10-JUN-1998; 98US-0088734P.
XX	10-JUN-1998; 98US-0088738P.
XX	10-JUN-1998; 98US-0088742P.
XX	10-JUN-1998; 98US-0088810P.
XX	10-JUN-1998; 98US-0088824P.
XX	10-JUN-1998; 98US-0088826P.
XX	11-JUN-1998; 98US-0088858P.
XX	11-JUN-1998; 98US-0088861P.
XX	11-JUN-1998; 98US-0088876P.
XX	12-JUN-1998; 98US-0089105P.
XX	16-JUN-1998; 98US-0089440P.
XX	16-JUN-1998; 98US-0089512P.
XX	16-JUN-1998; 98US-0089514P.
XX	17-JUN-1998; 98US-0089532P.
XX	17-JUN-1998; 98US-0089538P.
XX	17-JUN-1998; 98US-0089598P.
XX	17-JUN-1998; 98US-0089599P.
XX	17-JUN-1998; 98US-0089600P.
XX	17-JUN-1998; 98US-0089653P.
XX	18-JUN-1998; 98US-0089801P.
XX	18-JUN-1998; 98US-0089907P.
XX	18-JUN-1998; 98US-0089908P.
XX	19-JUN-1998; 98US-0089947P.
XX	19-JUN-1998; 98US-0089948P.
XX	19-JUN-1998; 98US-0089952P.
XX	22-JUN-1998; 98US-0090246P.
XX	22-JUN-1998; 98US-0090252P.
XX	22-JUN-1998; 98US-0090254P.
XX	22-JUN-1998; 98US-0090282P.
XX	23-JUN-1998; 98US-0090349P.
XX	23-JUN-1998; 98US-0090355P.
XX	24-JUN-1998; 98US-0090429P.
XX	24-JUN-1998; 98US-0090431P.

23-AUG-2000; 2000WO-US023522.  
24-AUG-2000; 2000WO-US023328.  
08-NOV-2000; 2000WO-US030952.  
01-DEC-2000; 2000WO-US032678.  
28-FEB-2001; 2001WO-US006520.  
01-JUN-2001; 2001WO-US017800.  
20-JUN-2001; 2001WO-US019692.  
29-JUN-2001; 2001WO-US021066.  
09-JUL-2001; 2001WO-US021735.  
28-AUG-2001; 2001US-00941992.  
(GETH ) GENENTECH INC.  
Ashkenazi AJ, Baker KP, Botstein D, Deenoyers L, Eaton DJ;  
Ferrara N, Fong S, Gerber H, Geritsen ME, Goddard A, Godowski PJ;  
Grimaldi JC, Gurney AL, Kijavini LJ, Napier MA, Pan J, Paoni NF;  
Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;  
Zhang Z;  
WPI; 2003-155950/15.  
New secreted and transmembrane PRO polypeptides (e.g. PRO183, PRO184,  
PRO361 or PRO846) useful as targets for therapeutic intervention in  
cancers (e.g. lung or breast cancers), or for diagnosing these cancers.  
Claim 12; Fig 290; 647pp; English.  
The invention discloses isolated PRO secreted/transmembrane polypeptides  
comprising a sequence without signal peptide and the nucleic acid  
encoding them. The polypeptides can be used to raise antibodies that  
specifically bind to the PRO polypeptide, for linking a bioactive  
molecule to a cell expressing a PRO protein and for modulating at least  
one biological activity of a cell. The PRO polypeptides or  
polynucleotides are also useful as pharmaceuticals, diagnostics,  
biosensors or bioreactors, for detecting or treating e.g. tumours in  
mammals, e.g. humans, dogs, cats, cattle, horses, sheep, goats or  
rabbits as targets for therapeutic intervention in certain cancers (e.g.  
colon, lung or breast cancers) and diagnostic determination of the  
presence of these cancers. The PRO polypeptides are also useful as  
molecular weight markers or for chromosome identification. The PRO genes  
are useful as hybridisation probes or for screening libraries of human  
cDNA, genomic DNA or mRNA. The PRO genes may also be used in gene  
therapy, particularly for replacing a defective gene. The sequences  
presented in ABUS8900-ABUS9046 are the PRO polypeptides of the invention  
Sequence 104 AA;  
Query Match 100.0%; Score 502; DB 6; Length 104;  
Best Local Similarity 100.0%; Pred. NO. 1.4e-48;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKLAALLGLCVALSCTSSAAAFVLSGSAKPVQAPVALESAAEAGACTLANPLGTLPKLL 60  
DB 1 MKLAALLGLCVALSCTSSAAAFVLSGSAKPVQAPVALESAAEAGACTLANPLGTLPKLL 60  
QY 61 LSSLGIPVNHLEGSQKVAELGQAVGAVKALKALIGALTTFVG 104  
DB 61 LSSLGIPVNHLEGSQKVAELGQAVGAVKALKALIGALTTFVG 104  
RESULT 16  
ABU92404  
ID ABU92404 standard; protein; 104 AA.  
XX ABU92404;  
AC ABU92404;  
XX 16-JUL-2003 (first entry)  
DT Novel human secreted and transmembrane protein PRO1245.  
DE Human; secreted and transmembrane protein; PRO; PRO183; PRO184; PRO185;  
KW PRO343; PRO1133; PRO331; PRO337; PRO363; PRO5723; PRO1114; PRO3301;  
KW PRO9940; PRO1181; PRO1170; PRO361; PRO846; bioactive molecule; toxin;

PR	24-JUN-1998;	98US-0090433P.
PR	24-JUN-1998;	98US-0090444P.
PR	24-JUN-1998;	98US-0090455P.
PR	24-JUN-1998;	98US-0090472P.
PR	24-JUN-1998;	98US-0090535P.
PR	24-JUN-1998;	98US-0090540P.
PR	24-JUN-1998;	98US-0090542P.
PR	24-JUN-1998;	98US-0090557P.
PR	24-JUN-1998;	98US-0090678P.
PR	25-JUN-1998;	98US-0090678P.
PR	25-JUN-1998;	98US-0090690P.
PR	25-JUN-1998;	98US-0090694P.
PR	25-JUN-1998;	98US-0090695P.
PR	25-JUN-1998;	98US-0090696P.
PR	26-JUN-1998;	98US-0090862P.
PR	26-JUN-1998;	98US-0090863P.
PR	01-JUL-1998;	98US-0091360P.
PR	01-JUL-1998;	98US-0091544P.
PR	02-JUL-1998;	98US-0091478P.
PR	02-JUL-1998;	98US-0091519P.
PR	02-JUL-1998;	98US-0091626P.
PR	02-JUL-1998;	98US-0091628P.
PR	02-JUL-1998;	98US-0091633P.
PR	02-JUL-1998;	98US-0091646P.
PR	02-JUL-1998;	98US-0091673P.
PR	07-JUL-1998;	98US-0091978P.
PR	07-JUL-1998;	98US-0091982P.
PR	09-JUL-1998;	98US-0092182P.
PR	10-JUL-1998;	98US-0092472P.
PR	20-JUL-1998;	98US-0093339P.
PR	30-JUL-1998;	98US-0094551P.
PR	04-AUG-1998;	98US-0095285P.
PR	04-AUG-1998;	98US-0095301P.
PR	04-AUG-1998;	98US-0095302P.
PR	04-AUG-1998;	98US-0095318P.
PR	04-AUG-1998;	98US-0095321P.
PR	04-AUG-1998;	98US-0095325P.
PR	10-AUG-1998;	98US-0095316P.
PR	10-AUG-1998;	98US-0095329P.
PR	10-AUG-1998;	98US-0096012P.
PR	11-AUG-1998;	98US-0096143P.
PR	11-AUG-1998;	98US-0096146P.
PR	12-AUG-1998;	98US-0096329P.
PR	17-AUG-1998;	98US-0096757P.
PR	17-AUG-1998;	98US-0096766P.
PR	17-AUG-1998;	98US-0096768P.
PR	17-AUG-1998;	98US-0096773P.
PR	17-AUG-1998;	98US-0096791P.
PR	17-AUG-1998;	98US-0096867P.
PR	17-AUG-1998;	98US-0096891P.
PR	17-AUG-1998;	98US-0096894P.
PR	17-AUG-1998;	98US-0096895P.
PR	17-AUG-1998;	98US-0096897P.
PR	18-AUG-1998;	98US-0096949P.
PR	18-AUG-1998;	98US-0096950P.
PR	18-AUG-1998;	98US-0096959P.
PR	18-AUG-1998;	98US-0096960P.
PR	18-AUG-1998;	98US-0097022P.
PR	19-AUG-1998;	98US-0097141P.
PR	20-AUG-1998;	98US-0097218P.
PR	24-AUG-1998;	98US-0097661P.
PR	26-AUG-1998;	98US-0097952P.
PR	26-AUG-1998;	98US-0097954P.
PR	26-AUG-1998;	98US-0097955P.
PR	26-AUG-1998;	98US-0097971P.
PR	26-AUG-1998;	98US-0097974P.
PR	26-AUG-1998;	98US-0097978P.
PR	26-AUG-1998;	98US-0097979P.
PR	26-AUG-1998;	98US-0097986P.
PR	26-AUG-1998;	98US-0098014P.
PR	31-AUG-1998;	98US-0098525P.
PR	16-SEP-1998;	98US-0100634P.
PR	16-SEP-1998;	98WO-US019330.
PR	17-SEP-1998;	98US-0100858P.
PR	17-SEP-1998;	98WO-US019437.
PR	07-OCT-1998;	98WO-US021141.
PR	01-DEC-1998;	98WO-US025108.
PR	22-DEC-1998;	98US-0113296P.
PR	05-JAN-1999;	99WO-US000106.
PR	20-FEB-1999;	99WO-US030911.
PR	08-MAR-1999;	99WO-US005028.
PR	12-MAR-1999;	99US-0123957P.
PR	02-JUN-1999;	99WO-US012252.
PR	23-JUN-1999;	99US-0141037P.
PR	07-JUL-1999;	99US-0143048P.
PR	20-JUL-1999;	99US-0144758P.
PR	26-JUL-1999;	99US-0145698P.
PR	28-JUL-1999;	99US-0146232P.
PR	17-AUG-1999;	99US-0149396P.
PR	15-SEP-1999;	99WO-US021090.
PR	15-SEP-1999;	99WO-US021547.
PR	08-OCT-1999;	99US-0158663P.
PR	30-NOV-1999;	99WO-US028313.
PR	01-DEC-1999;	99WO-US028301.
PR	01-DEC-1999;	99WO-US028634.
PR	16-DEC-1999;	99WO-US030095.
PR	05-JAN-2000;	2000WO-US000219.
PR	06-JAN-2000;	2000WO-US000376.
PR	11-FEB-2000;	2000WO-US003565.
PR	18-FEB-2000;	2000WO-US004341.
PR	22-FEB-2000;	2000WO-US004414.
PR	24-FEB-2000;	2000WO-US004914.
PR	02-MAR-2000;	2000WO-US005004.
PR	02-MAR-2000;	2000WO-US005841.
PR	10-MAR-2000;	2000WO-US006319.
PR	15-MAR-2000;	2000WO-US006884.
PR	20-MAR-2000;	2000WO-US007377.
PR	30-MAR-2000;	2000WO-US008439.
PR	15-MAY-2000;	2000WO-US013358.
PR	17-MAY-2000;	2000WO-US013705.
PR	22-MAY-2000;	2000WO-US014042.
PR	30-MAY-2000;	2000WO-US014941.
PR	02-JUN-2000;	2000WO-US015264.
PR	23-JUN-2000;	2000US-0213637P.
PR	28-JUL-2000;	2000WO-US020710.
PR	11-AUG-2000;	2000WO-US022031.
Query Match 100.0%; Score 502; DB 6; Length 104;		
Best Local Similarity 100.0%; Pred. No. 1.4e-48;		
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	MKLAALLGLCVALSCTSSAAAFVLSAKFVAQFVAALSSAAEAGAGTLANPLGTINPLKLL 60
Db	1	MKLAALLGLCVALSCTSSAAAFVLSAKFVAQFVAALSSAAEAGAGTLANPLGTINPLKLL 60
Qy	61	LSLGIPTVNNHLEGSQKCVAEELGPOAVGAVKALKALLGALTIVFG 104
Db	61	LSLGIPTVNNHLEGSQKCVAEELGPOAVGAVKALKALLGALTIVFG 104
RESULT 17		
ABUS9469		
ID	ABUS9469 standard; protein; 104 AA.	
XX	ABUS9469;	
AC		
XX	22-APR-2003 (first entry)	
DT	Novel human secreted or transmembrane protein PRO1358.	
DE	Human, PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;	
KW	cardiac insufficiency disorder; cancer; tumour; immune response;	
KW	adrenal cortical capillary endothelial growth; c-fos induction;	
KW	vascular endothelial growth factor inhibition; VEGF inhibition;	
KW	endothelial cell growth inhibitor; T-lymphocytes stimulation;	
KW	retinal neurons cell survival; rod photoreceptor cell survival;	

retinal disorder; retinitis pigmentosa; kidney disorder;  
mammalian kidney mesangial cell proliferation; Berger disease;  
maternal; herpetiformis; Crohn's disease; chondrocyte proliferation;  
chondrocyte redifferentiation; sports injury; arthritis.

KW	retinal disorder; retinitis pigmentosa; kidney disorder;	PR	24-JUN-1998;	98US-0090435P.
KW	mammalian kidney mesangial cell proliferation; Berger disease;	PR	24-JUN-1998;	98US-0090444P.
KW	maternal; herpetiformis; Crohn's disease; chondrocyte proliferation;	PR	24-JUN-1998;	98US-0090445P.
KW	chondrocyte redifferentiation; sports injury; arthritis.	PR	24-JUN-1998;	98US-0090535P.
XX		PR	24-JUN-1998;	98US-0090540P.
OS	Homo sapiens.	PR	24-JUN-1998;	98US-0090542P.
XX		PR	24-JUN-1998;	98US-0090557P.
PN	US2003027985-A1.	PR	24-JUN-1998;	98US-0090676P.
XX		PR	25-JUN-1998;	98US-0090678P.
PD		PR	25-JUN-1998;	98US-0090890P.
XX		PR	25-JUN-1998;	98US-0090894P.
PF		PR	25-JUN-1998;	98US-0090695P.
XX		PR	25-JUN-1998;	98US-0090696P.
XX		PR	25-JUN-1998;	98US-0090696P.
PR		PR	26-JUN-1998;	98US-0090862P.
PR		PR	26-JUN-1998;	98US-0090863P.
PR		PR	01-JUL-1998;	98US-0091360P.
PR		PR	01-JUL-1998;	98US-0091544P.
PR		PR	02-JUL-1998;	98US-0091478P.
PR		PR	02-JUL-1998;	98US-0091519P.
PR		PR	02-JUL-1998;	98US-0091626P.
PR		PR	02-JUL-1998;	98US-0091628P.
PR		PR	02-JUL-1998;	98US-0091633P.
PR		PR	02-JUL-1998;	98US-0091646P.
PR		PR	02-JUL-1998;	98US-0091673P.
PR		PR	07-JUL-1998;	98US-0091978P.
PR		PR	07-JUL-1998;	98US-0091982P.
PR		PR	09-JUL-1998;	98US-0092182P.
PR		PR	10-JUL-1998;	98US-0092472P.
PR		PR	20-JUL-1998;	98US-0093339P.
PR		PR	30-JUL-1998;	98US-0094651P.
PR		PR	04-AUG-1998;	98US-0095282P.
PR		PR	04-AUG-1998;	98US-0095285P.
PR		PR	04-AUG-1998;	98US-0095301P.
PR		PR	04-AUG-1998;	98US-0095318P.
PR		PR	04-AUG-1998;	98US-0095321P.
PR		PR	04-AUG-1998;	98US-0095325P.
PR		PR	10-AUG-1998;	98US-0095316P.
PR		PR	10-AUG-1998;	98US-0095929P.
PR		PR	10-AUG-1998;	98US-0096012P.
PR		PR	11-AUG-1998;	98US-0096143P.
PR		PR	11-AUG-1998;	98US-0096146P.
PR		PR	12-AUG-1998;	98US-0096329P.
PR		PR	17-AUG-1998;	98US-0096757P.
PR		PR	17-AUG-1998;	98US-0096766P.
PR		PR	17-AUG-1998;	98US-0096768P.
PR		PR	17-AUG-1998;	98US-0096773P.
PR		PR	17-AUG-1998;	98US-0096791P.
PR		PR	17-AUG-1998;	98US-0096867P.
PR		PR	17-AUG-1998;	98US-0096891P.
PR		PR	17-AUG-1998;	98US-0096894P.
PR		PR	17-AUG-1998;	98US-0096895P.
PR		PR	17-AUG-1998;	98US-0096897P.
PR		PR	18-AUG-1998;	98US-0096949P.
PR		PR	18-AUG-1998;	98US-0096950P.
PR		PR	18-AUG-1998;	98US-0096959P.
PR		PR	18-AUG-1998;	98US-0096960P.
PR		PR	18-AUG-1998;	98US-0097022P.
PR		PR	19-AUG-1998;	98US-0097218P.
PR		PR	20-AUG-1998;	98US-0097661P.
PR		PR	26-AUG-1998;	98US-0097952P.
PR		PR	26-AUG-1998;	98US-0097954P.
PR		PR	26-AUG-1998;	98US-0097955P.
PR		PR	26-AUG-1998;	98US-0097971P.
PR		PR	26-AUG-1998;	98US-0097974P.
PR		PR	26-AUG-1998;	98US-0097978P.
PR		PR	26-AUG-1998;	98US-0097979P.
PR		PR	26-AUG-1998;	98US-0097986P.
PR		PR	31-AUG-1998;	98US-0098014P.
PR		PR	31-AUG-1998;	98US-0098525P.
PR		PR	16-SEP-1998;	98US-0100634P.

[illegible]

XX	ABU10941;	
XX	AC	
XX	XX	
XX	DT	
XX	XX	
XX	DE	
XX	XX	
XX	Human PRO polypeptide #127.	
XX	Human; PRO; secreted polypeptide; transmembrane polypeptide; toxin;	
XX	radiolabel; cell death; gene mapping; chromosome mapping;	
XX	protein electrophoresis; genetic disorder; immunosuppressive; cytostatic;	
XX	antibacterial.	
XX	XX	
XX	XX	
OS	Homo sapiens.	
XX	XX	
PN	US2002123463-A1.	
XX	XX	
XX	05-SEP-2002.	
PD	XX	
XX	XX	
PF	19-NOV-2001; 2001US-00989732.	

```

XX 16-JUN-1997; 97US-0049787P.
PR 17-OCT-1997; 97US-0062250P.
PR 05-NOV-1997; 97US-0062250P.
PR 12-NOV-1997; 97US-0065186P.
PR 13-NOV-1997; 97US-0065311P.
PR 24-NOV-1997; 97US-0066770P.
PR 25-FEB-1998; 98US-0075945P.
PR 20-MAR-1998; 98US-0078910P.
PR 28-APR-1998; 98US-0083322P.
PR 27-MAY-1998; 98US-0084600P.
PR 08-JUN-1998; 98US-0087108P.
PR 02-JUN-1998; 98US-0087607P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088021P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088030P.
PR 04-JUN-1998; 98US-0088033P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088214P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088734P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088742P.
PR 10-JUN-1998; 98US-0088810P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088858P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089440P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089532P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089599P.
PR 17-JUN-1998; 98US-0089600P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089601P.
PR 18-JUN-1998; 98US-0089907P.
PR 18-JUN-1998; 98US-0089908P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98WO-US019330.
PR 07-OCT-1998; 98WO-US021141.
PR 01-DEC-1998; 98WO-US025108.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 02-JUN-1999; 99WO-US012252.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 16-DEC-1999; 99WO-US030095.
PR 20-DEC-1999; 99WO-US030911.
PR 06-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 23-FEB-2000; 2000WO-US004414.
PR 24-FEB-2000; 2000WO-US004914.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.

PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 30-MAR-2000; 2000WO-US008439.
PR 15-MAY-2000; 2000WO-US013358.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023328.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030352.
PR 01-DEC-2000; 2000WO-US032578.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-JUN-2001; 2001WO-US017800.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 28-AUG-2001; 2001US-00941992.

XX (GETH ) GENENTECH INC.
PA Ashkenazi AJ, Baker KF, Botstein D, Desnoyers L, Eaton DL,
XX Ferrara N, Fong S, Garber H, Gerritsen ME, Goddard A, Godowski PJ,
PI Grimaldi JC, Gurney AL, Klijavin LJ, Napier MA, Pan J, Paoni NF,
PI Roy MA, Stewart TA, Tumas D, Watanabe CX, Williams PM, Wood WI,
PI Zhang Z;
XX WPI; 2003-066810/06.
DR N-PSDB; ABX17173.
XX
DR Novel secreted and transmembrane polypeptide for modulating biological
PT activity of cell expressing the polypeptide, identifying agonists or
PT antagonists of polypeptide, and as molecular weight markers.
XX
XX Claim 12; Fig 290; 655pp; English.
XX
XX The invention relates to a secreted and transmembrane polypeptide, termed
CC PRO polypeptide, and the polynucleotide encoding it. The polypeptide is
CC useful for detecting PRO polypeptides and for linking a bioactive
CC molecule to a cell expressing the above polypeptides, where the bioactive
CC molecule is a toxin, radiolabel or an antibody. The bioactive material
CC causes the death of the cell. The polypeptide is useful for identifying
CC agonists or antagonists of the PRO polypeptide, for preparing variants of
CC PRO, as a molecular weight marker for protein electrophoresis purposes
CC and the PRO polynucleotide is useful for recombinantly expressing those
CC markers. The polynucleotide is also useful as a hybridisation probe, in
CC chromosome and gene mapping, in generation of antisense RNA and DNA, in
CC the preparation of PRO polypeptide, for generating transgenic animals or
CC knockout animals which in turn are useful in the development and
CC screening of therapeutically useful reagents, to construct hybridisation
CC probes for mapping the gene which encodes PRO and for the genetic
CC analysis of individuals with genetic disorders, in gene therapy, for
CC chromosome identification, as a chromosome marker and for generating
CC probes for PCR, Northern analysis, Southern analysis and Western
CC analysis. This sequence represents a human PRO polypeptide of the
XX invention
XX
XX Sequence 104 AA;
XX
Query Watch 100.0%; Score 502; DB 6; Length 104;
Best Local Similarity 100.0%; Pred. NO. 1.4e-48;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLAALLGLCVALLSCSSAAAFVLSAKFVAQPVAALESAAEAGAGTLANPLGTLNPKLL 60
Db 1 MKLAALLGLCVALLSCSSAAAFVLSAKFVAQPVAALESAAEAGAGTLANPLGTLNPKLL 60
QY 61 LSSLGIPVNHLEGSQKCVAEELGPOAVGAVKALKALIGALTIVFG 104
Db 61 LSSLGIPVNHLEGSQKCVAEELGPOAVGAVKALKALIGALTIVFG 104

```



CC useful as a therapeutic agent e.g. for treating cancer and autoimmune  
CC disease. PRO is useful in assays to identify other proteins or molecules  
CC involved in binding interactions. The polynucleotide (ii) encoding (i) is  
CC useful in chromosome and gene mapping, for generating transgenic animals  
CC or knockout animals which in turn are useful in the development and  
CC screening of therapeutically useful reagents, for the genetic analysis of  
CC individuals with genetic disorders, in gene therapy, for chromosome  
CC identification, and as a chromosome marker. An anti-(i)-antibody is  
CC useful in diagnostic assays for PRO, e.g. detecting its expression in  
CC specific cells, tissues or serum, for affinity purification of PRO, and  
CC for treating septic shock. This is the amino acid sequence of a novel  
CC human secreted and transmembrane PRO polypeptide  
XX  
SQ

Query Match 100.0%; Score 502; DB 6; Length 104;  
Best Local Similarity 100.0%; Pred. No. 1.4e-48;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKLAALIGLCVALSCSSAAATLVGSAKPVQAVPVALESAREAGATLANPLGTINPKLL 60  
Db 1 MKLAALIGLCVALSCSSAAATLVGSAKPVQAVPVALESAREAGATLANPLGTINPKLL 60  
QY 61 LSSLGIPVNHILIEGSKVVAELGPOAVGAVKALKALIGALTIVFG 104  
Db 61 LSSLGIPVNHILIEGSKVVAELGPOAVGAVKALKALIGALTIVFG 104

RESULT 21  
ABU88632  
ID ABU88632 standard; protein; 104 AA.  
AC ABU88632;  
XX  
XX  
DT 11-AUG-2003 (first entry)  
XX  
DE Human secreted and transmembrane polypeptide PRO1245.  
DE  
DE Human; gene therapy; cancer; retinal disorder; wound healing;  
KW kidney disorder.  
KW  
KW Homo sapiens.  
XX  
XX  
OS US2002197615-A1.  
PN  
XX  
XX  
PD 26-DEC-2002.  
XX  
XX  
PF 16-NOV-2001; 2001US-00991181.  
XX

XX 16-JUN-1997; 97US-0049787P.  
XX 17-OCT-1997; 97US-0062250P.  
XX 05-NOV-1997; 97WO-US020069.  
XX 12-NOV-1997; 97US-0065186P.  
XX 13-NOV-1997; 97US-0065311P.  
XX 24-NOV-1997; 97US-0066770P.  
XX 23-FEB-1998; 98US-0075945P.  
XX 20-MAR-1998; 98US-0078910P.  
XX 28-APR-1998; 98US-0083322P.  
XX 07-MAY-1998; 98US-0084600P.  
XX 28-MAY-1998; 98US-0087106P.  
XX 02-JUN-1998; 98US-0087607P.  
XX 02-JUN-1998; 98US-0087609P.  
XX 02-JUN-1998; 98US-0087759P.  
XX 03-JUN-1998; 98US-0087827P.  
XX 04-JUN-1998; 98US-0088021P.  
XX 04-JUN-1998; 98US-0088025P.  
XX 04-JUN-1998; 98US-0088026P.  
XX 04-JUN-1998; 98US-0088028P.  
XX 04-JUN-1998; 98US-0088029P.  
XX 04-JUN-1998; 98US-0088030P.  
XX 04-JUN-1998; 98US-0088033P.  
XX 04-JUN-1998; 98US-0088326P.  
XX 05-JUN-1998; 98US-0088167P.

PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088555P.  
PR 10-JUN-1998; 98US-0088734P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088742P.  
PR 10-JUN-1998; 98US-0088810P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088858P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089440P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089532P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089599P.  
PR 17-JUN-1998; 98US-0089600P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089801P.  
PR 18-JUN-1998; 98US-0089907P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 98WO-US000106.  
PR 08-MAR-1999; 98WO-US005028.  
PR 02-JUN-1999; 98WO-US012252.  
PR 15-SEP-1999; 98WO-US021090.  
PR 15-SEP-1999; 98WO-US021547.  
PR 30-NOV-1999; 98WO-US028313.  
PR 01-DEC-1999; 98WO-US028301.  
PR 01-DEC-1999; 98WO-US028634.  
PR 16-DEC-1999; 98WO-US030095.  
PR 20-DEC-1999; 98WO-US030911.  
PR 05-JAN-2000; 2000WO-US000319.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 15-MAY-2000; 2000WO-US013358.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 28-AUG-2001; 2001US-00941992.  
XX  
XX  
XX (GETH ) GENENTECH INC.

Ashtkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;



PI	Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;	
PI	Grimaldi JC, Gurney AL, Kijavini IG, Napier MA, Pan J, Paoni NF;	
PI	Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;	
PI	Zhang Z;	
XX	WPI: 2003-370792/35.	
DR	N-PSDB; ACA88477.	
DR		
XX		
PT	New secreted and transmembrane nucleic acids and polypeptides, designated	
PT	as PRO, useful for the preparation of a medicament for treating a	
PT	condition that is responsive to the PRO polypeptide. e.g., cancer.	
XX		
PS	Claim 12; Fig 290; 647pp; English.	
XX		
CC	The invention relates to an isolated nucleic acid encoding a PRO	
CC	polypeptide. The polypeptide, agonist, antagonist and antibody are useful	
CC	for the preparation of a medicament for treating a condition that is	
CC	responsive to the PRO polypeptide. The nucleotide sequence is useful in	
CC	molecular biology including being used as hybridisation probes, in	
CC	chromosome and gene mapping and in the generation of anti-sense RNA and	
CC	DNA. The PRO polypeptides can also be used in the treatment of e.g.	
CC	cancer, retinal disorders, wound healing and kidney disorders. The	
CC	present sequence represents the amino acid sequence of a human secreted	
CC	and transmembrane PRO polypeptide of the present invention. Note: The	
CC	sequence data for this patent did not form part of the printed	
CC	specification but was obtained in electronic format directly from USPTO	
CC	at seqdata.uspto.gov/sequence.html?DocID=20020197615	
XX		
SQ	Sequence 104 AA;	
	Query Match 100.0%; Score 502; DB 6; Length 104;	
	Best Local Similarity 100.0%; Pred. No. 1.4e-48;	
	Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	1 MKLAALGLCVALSCESSAAFLVGSAPVAPVVALESAAAGAGTLANPLGLINPKLL 60	
DB	1 MKLAALGLCVALSCESSAAFLVGSAPVAPVVALESAAAGAGTLANPLGLINPKLL 60	
QY	61 LSSLGIPVNHLEIGSKCKVVAELGPOAVGAVKALKALLGALT VFG 104	
DB	61 LSSLGIPVNHLEIGSKCKVVAELGPOAVGAVKALKALLGALT VFG 104	
RESULT 22		
ABO34146		
ID	ABO34146 standard; protein; 104 AA.	
XX		
AC	ABO34146;	
XX		
DT	19-SEP-2003 (first entry)	
XX		
DE	Human PRO1245 polypeptide.	
XX		
KW	Human; PRO polypeptide; secreted protein; transmembrane protein;	
KW	biosensor; bioreactor; tumour; cancer; diabetes; ALS; ulcer;	
KW	rheumatoid arthritis; amyotrophic lateral sclerosis; cytostatic;	
KW	antidiabetic; antiarthritic; antirheumatic; antiulcer.	
XX		
OS	Homo sapiens.	
XX		
PN	US2003017981-A1.	
XX		
FD	23-JAN-2003.	
XX		
PF	20-NOV-2001; 2001US-00989728.	
XX		
FR	16-JUN-1997; 97US-0049787P.	
PR	17-OCT-1997; 97US-0062250P.	
PR	05-NOV-1997; 97WO-US020069.	
PR	12-NOV-1997; 97US-0065186P.	
PR	13-NOV-1997; 97US-0065531P.	
PR	24-NOV-1997; 97US-0066770P.	
PR	25-FEB-1998; 98US-0075945P.	

PR	02-JUL-1998;	98US-0091626P.	PR	05-JAN-2000;	200WO-US000219.
PR	02-JUL-1998;	98US-0091633P.	PR	06-JAN-2000;	200WO-US000376.
PR	07-JUL-1998;	98US-0091582P.	PR	11-FEB-2000;	200WO-US003565.
PR	08-JUL-1998;	98US-0092182P.	PR	18-FEB-2000;	200WO-US004341.
PR	10-JUL-1998;	98US-0092472P.	PR	22-FEB-2000;	200WO-US004414.
PR	20-JUL-1998;	98US-0093339P.	PR	24-FEB-2000;	200WO-US004914.
PR	30-JUL-1998;	98US-0094651P.	PR	24-FEB-2000;	200WO-US005004.
PR	04-AUG-1998;	98US-0095282P.	PR	02-MAR-2000;	200WO-US005841.
PR	04-AUG-1998;	98US-0095301P.	PR	10-MAR-2000;	200WO-US006319.
PR	04-AUG-1998;	98US-0095302P.	PR	15-MAR-2000;	200WO-US006884.
PR	04-AUG-1998;	98US-0095318P.	PR	20-MAR-2000;	200WO-US007377.
PR	04-AUG-1998;	98US-0095321P.	PR	30-MAR-2000;	200WO-US008439.
PR	04-AUG-1998;	98US-0095325P.	PR	15-MAY-2000;	200WO-US013358.
PR	10-AUG-1998;	98US-0095929P.	PR	17-MAY-2000;	200WO-US013705.
PR	10-AUG-1998;	98US-0096012P.	PR	22-MAY-2000;	200WO-US014042.
PR	11-AUG-1998;	98US-0096143P.	PR	30-MAY-2000;	200WO-US014941.
PR	11-AUG-1998;	98US-0096146P.	PR	02-JUN-2000;	200WO-US015264.
PR	12-AUG-1998;	98US-0096329P.	PR	23-JUN-2000;	200WO-US013637P.
PR	17-AUG-1998;	98US-0096757P.	PR	28-JUL-2000;	200WO-US020710.
PR	17-AUG-1998;	98US-0096766P.	PR	11-AUG-2000;	200WO-US020231.
PR	17-AUG-1998;	98US-0096773P.	PR	23-AUG-2000;	200WO-US023522.
PR	17-AUG-1998;	98US-0096791P.	PR	24-AUG-2000;	200WO-US023328.
PR	17-AUG-1998;	98US-0096867P.	PR	07-SEP-2000;	200WO-US023078P.
PR	17-AUG-1998;	98US-0096881P.	PR	08-NOV-2000;	200WO-US030952.
PR	17-AUG-1998;	98US-0096894P.	PR	01-DEC-2000;	200WO-US032678.
PR	17-AUG-1998;	98US-0096897P.	PR	28-FEB-2001;	200WO-US006520.
PR	18-AUG-1998;	98US-0096950P.			
PR	18-AUG-1998;	98US-0096959P.			
PR	18-AUG-1998;	98US-0096960P.			
PR	18-AUG-1998;	98US-0097022P.			
PR	18-AUG-1998;	98US-0097141P.			
PR	20-AUG-1998;	98US-0097218P.			
PR	24-AUG-1998;	98US-0097661P.			
PR	26-AUG-1998;	98US-0097952P.			
PR	26-AUG-1998;	98US-0097954P.			
PR	26-AUG-1998;	98US-0097955P.			
PR	26-AUG-1998;	98US-0097971P.			
PR	26-AUG-1998;	98US-0097974P.			
PR	26-AUG-1998;	98US-0097978P.			
PR	26-AUG-1998;	98US-0097979P.			
PR	26-AUG-1998;	98US-0097986P.			
PR	26-AUG-1998;	98US-0098014P.			
PR	31-AUG-1998;	98US-0098525P.			
PR	16-SEP-1998;	98US-0100634P.			
PR	17-SEP-1998;	98US-0100858P.			
PR	17-SEP-1998;	98US-0100858P.			
PR	07-OCT-1998;	98US-01019437.			
PR	01-DEC-1998;	98US-01021141.			
PR	22-DEC-1998;	98US-0113296P.			
PR	05-JAN-1999;	99WO-US000106.			
PR	08-MAR-1999;	99WO-US005028.			
PR	12-MAR-1999;	99US-0123957P.			
PR	23-JUN-1999;	99WO-US012252.			
PR	07-JUL-1999;	99US-0141037P.			
PR	20-JUL-1999;	99US-0144758P.			
PR	26-JUL-1999;	99US-0145698P.			
PR	28-JUL-1999;	99US-0146222P.			
PR	17-AUG-1999;	99US-0149396P.			
PR	15-SEP-1999;	99WO-US021090.			
PR	15-SEP-1999;	99WO-US021547.			
PR	08-OCT-1999;	99US-0158663P.			
PR	30-NOV-1999;	99WO-US028313.			
PR	01-DEC-1999;	99WO-US028301.			
PR	01-DEC-1999;	99WO-US028634.			
PR	16-DEC-1999;	99WO-US030095.			
PR	20-DEC-1999;	99WO-US030911.			
Query Match 100.0%; Score 502; DB 6; Length 104;					
Best Local Similarity 100.0%; Pred. No. 1.4e-48; Mismatches 0; Indels 0; Gaps 0;					
Matches 104; Conservative 0;					
QY	1	MKLAALLGLCVALS	SSAAAF	LVGSAKPVAQ	PVAAL
Db	1	MKLAALLGLCVALS	SSAAAF	LVGSAKPVAQ	PVAAL
QY	61	LSSLGIPVNHLE	EGSKCV	AEILGPOAV	GAVKALKALGALT
Db	61	LSSLGIPVNHLE	EGSKCV	AEILGPOAV	GAVKALKALGALT
RESULT 23					
ADA37919					
ID	ADA37919 standard; protein; 104 AA.				
XX	ADA37919;				
XX	20-NOV-2003 (first entry)				
XX	Human secreted/transmembrane protein PRO1245.				
XX	PRO; secreted protein; transmembrane protein;				
KW	hypertrophy of neonatal heart; angiogenesis;				
KW	vascular endothelial growth factor; VEGF-stimulated proliferation;				
KW	endothelial cell; T-lymphocyte proliferation; retinal neuron;				
KW	c-fos induction; adipocyte cell; chondrocyte differentiation;				
KW	pancreatic beta-cell precursor differentiation; Gene therapy; tumour;				
KW	cancer; human; colon cancer; lung cancer; breast cancer;				
KW	rod photoreceptor cell.				
OS	Homo sapiens.				
XX	US2003008297-A1.				
XX	09-JAN-2003.				
XX	15-NOV-2001; 2001US-00997653.				
XX	16-JUN-1997; 97US-0049787P.				
PR	17-OCT-1997; 97US-0062250P.				
PR	05-NOV-1997; 97WO-US020069.				
PR	12-NOV-1997; 97US-0065186P.				
PR	13-NOV-1997; 97US-0065511P.				
PR	24-NOV-1997; 97US-0065677Q.				

PR 25-FEB-1998; 98US-0075945P.  
 PR 28-MAR-1998; 98US-0078910P.  
 PR 28-APR-1998; 98US-0083322P.  
 PR 28-MAY-1998; 98US-0084600P.  
 PR 28-MAY-1998; 98US-0087106P.  
 PR 02-JUN-1998; 98US-0087607P.  
 PR 02-JUN-1998; 98US-0087609P.  
 PR 02-JUN-1998; 98US-0087759P.  
 PR 03-JUN-1998; 98US-0088021P.  
 PR 04-JUN-1998; 98US-0088025P.  
 PR 04-JUN-1998; 98US-0088026P.  
 PR 04-JUN-1998; 98US-0088028P.  
 PR 04-JUN-1998; 98US-0088029P.  
 PR 04-JUN-1998; 98US-0088030P.  
 PR 04-JUN-1998; 98US-0088033P.  
 PR 05-JUN-1998; 98US-0088167P.  
 PR 05-JUN-1998; 98US-0088202P.  
 PR 05-JUN-1998; 98US-0088212P.  
 PR 05-JUN-1998; 98US-0088217P.  
 PR 09-JUN-1998; 98US-0088655P.  
 PR 10-JUN-1998; 98US-0088734P.  
 PR 10-JUN-1998; 98US-0088738P.  
 PR 10-JUN-1998; 98US-0088742P.  
 PR 10-JUN-1998; 98US-0088810P.  
 PR 10-JUN-1998; 98US-0088824P.  
 PR 10-JUN-1998; 98US-0088826P.  
 PR 11-JUN-1998; 98US-0088858P.  
 PR 11-JUN-1998; 98US-0088861P.  
 PR 11-JUN-1998; 98US-0088876P.  
 PR 12-JUN-1998; 98US-0089105P.  
 PR 16-JUN-1998; 98US-0089440P.  
 PR 16-JUN-1998; 98US-0089512P.  
 PR 16-JUN-1998; 98US-0089514P.  
 PR 17-JUN-1998; 98US-0089532P.  
 PR 17-JUN-1998; 98US-0089538P.  
 PR 17-JUN-1998; 98US-0089598P.  
 PR 17-JUN-1998; 98US-0089599P.  
 PR 17-JUN-1998; 98US-0089600P.  
 PR 17-JUN-1998; 98US-0089653P.  
 PR 18-JUN-1998; 98US-0089801P.  
 PR 18-JUN-1998; 98US-0089907P.  
 PR 18-JUN-1998; 98US-0089908P.  
 PR 16-SEP-1998; 98WO-US019330.  
 PR 17-SEP-1998; 98WO-US019437.  
 PR 07-OCT-1998; 98WO-US021141.  
 PR 01-DEC-1998; 98WO-US025108.  
 PR 05-JAN-1999; 99WO-US000106.  
 PR 08-MAR-1999; 99WO-US005028.  
 PR 02-JUN-1999; 99WO-US012252.  
 PR 15-SEP-1999; 99WO-US021090.  
 PR 15-SEP-1999; 99WO-US021547.  
 PR 30-NOV-1999; 99WO-US028313.  
 PR 01-DEC-1999; 99WO-US028301.  
 PR 01-DEC-1999; 99WO-US028334.  
 PR 16-DEC-1999; 99WO-US030095.  
 PR 20-DEC-1999; 99WO-US030911.  
 PR 05-JAN-2000; 2000WO-US000219.  
 PR 06-JAN-2000; 2000WO-US000376.  
 PR 11-FEB-2000; 2000WO-US003565.  
 PR 18-FEB-2000; 2000WO-US004341.  
 PR 22-FEB-2000; 2000WO-US004414.  
 PR 24-FEB-2000; 2000WO-US004914.  
 PR 02-MAR-2000; 2000WO-US005004.  
 PR 02-MAR-2000; 2000WO-US005841.  
 PR 15-MAR-2000; 2000WO-US006319.  
 PR 15-MAR-2000; 2000WO-US006884.  
 PR 30-MAR-2000; 2000WO-US007377.  
 PR 30-MAR-2000; 2000WO-US008439.  
 PR 15-MAY-2000; 2000WO-US013358.  
 PR 17-MAY-2000; 2000WO-US013705.  
 PR 22-MAY-2000; 2000WO-US014042.

PR 30-MAY-2000; 2000WO-US014941.  
 PR 02-JUN-2000; 2000WO-US015264.  
 PR 28-JUL-2000; 2000WO-US020710.  
 PR 11-AUG-2000; 2000WO-US022031.  
 PR 23-AUG-2000; 2000WO-US023522.  
 PR 24-AUG-2000; 2000WO-US023328.  
 PR 08-NOV-2000; 2000WO-US030952.  
 PR 01-DEC-2000; 2000WO-US032678.  
 PR 28-FEB-2001; 2001WO-US006520.  
 PR 01-JUN-2001; 2001WO-US017800.  
 PR 20-JUN-2001; 2001WO-US019692.  
 PR 29-JUN-2001; 2001WO-US021066.  
 PR 09-JUL-2001; 2001WO-US021735.  
 PR 28-AUG-2001; 2001US-00941992.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Ashkenazi AJ, Baker KP, Botstein D, Deenoyers L, Raton DL,  
 PI Perrara N, Fong S, Gerber H, Geritsen ME, Goddard A, Godowski PJ,  
 PI Grimaldi JC, Gurney AL, Kijavini LJ, Napier MA, Pan J, Paoni NF,  
 PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI,  
 PI Zhang Z;  
 XX  
 DR WPI; 2003-531419/50.  
 DR N-PSDB; ADA37918.  
 XX  
 PT New isolated PRO183, PRO184, PRO361 or PRO846 nucleic acid and secreted  
 PT transmembrane polypeptides, useful as targets for the diagnosis and  
 PT treatment of cancers, such as lung and breast cancers.  
 XX  
 PS Claim 12; Fig 290; 660pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule comprising the  
 CC full-length coding sequence of the DNA ATCC Accession Numbers given in  
 CC the specification, or comprising a sequence with at least 80% identity  
 CC to: (a) a nucleotide encoding any of 147 PRO polypeptides, or an  
 CC extracellular domain of the polypeptide; or (b) any of 147 nucleotide  
 CC sequences fully defined in the specification. Also included are the PRO  
 CC proteins (or their extracellular domains with or without their associated  
 CC extracellular domains), expression vectors, host cells, PRO chimeric  
 CC proteins, anti-PRO antibodies, methods of detecting polypeptide in a  
 CC sample, methods of linking a bioactive molecule to a cell expressing a  
 CC polypeptide and methods of modulating at least one biological activity of  
 CC a cell expressing the polypeptide. The PRO polypeptides or  
 CC polynucleotides are useful as pharmaceuticals, diagnostics, or  
 CC bioreactors. These are useful for stimulating hypertrophy of neonatal  
 CC heart, promoting angiogenesis, inhibiting vascular endothelial growth  
 CC factor (VEGF)-stimulated proliferation of endothelial cells, modulating  
 CC the proliferation of stimulated T-lymphocytes, enhancing the survival or  
 CC proliferation of retinal neurons or rod photoreceptor cells, inducing c-  
 CC fos in endothelial cells, modulating glucose or FFA uptake by adipocyte  
 CC cells, inducing proliferation and/or re-differentiation of chondrocytes,  
 CC or inducing pancreatic beta-cell precursor differentiation. In  
 CC particular, these are useful for detecting or treating tumours and  
 CC certain cancers (colon, lung or breast cancers) in mammals, e.g. humans,  
 CC dogs, cats, cattle, horses, sheep, pigs, goats, or rabbits. The PRO genes  
 CC may also be used in gene therapy, particularly for replacing a defective  
 CC gene. The present sequence represents a PRO protein.  
 XX  
 SQ Sequence 104 AA;

Query Match 100.0%; Score 502; DB 6; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-48;  
 Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKLAALLGLCVALLSCSSAAFLVGSAPVQVPAALSAEAGAGTLANPLTLNPLKLL 60  
 DB 1 MKLAALLGLCVALLSCSSAAFLVGSAPVQVPAALSAEAGAGTLANPLTLNPLKLL 60  
 QY 61 LSSIGIPVNHIEGSKCVAEELGPQAVGAVKALKALLGALTVEG 104  
 DB 61 LSSIGIPVNHIEGSKCVAEELGPQAVGAVKALKALLGALTVEG 104

RESULT 24  
ADA21605  
ID ADA21605 standard; protein; 104 AA.  
XX AC ADA21605;  
XX AC ADA21605;  
XX AC ADA21605;  
DT DT 20-NOV-2003 (first entry)  
DE DE Human secreted/transmembrane polypeptide PRO1245.  
XX XX  
KW human; tumour; cancer; colorectal cancer; gene therapy;  
KW chondrocyte differentiation; VEGF inhibition;  
KW vascular endothelial growth factor; Alzheimer's disease;  
KW Parkinson's disease; atherosclerosis; cystic fibrosis;  
KW multiple sclerosis; ovarian cancer; tissue typing.  
XX XX  
CS Homo sapiens.  
XX XX  
XX US2003054404-A1.  
XX PD 20-MAR-2003.  
XX XX  
XX 15-NOV-2001; 2001US-00997601.  
XX 16-JUN-1997; 97US-0049787P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 05-NOV-1997; 97WO-US020069.  
PR 12-NOV-1997; 97US-0065186P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 24-NOV-1997; 97US-0066770P.  
PR 25-FEB-1998; 98US-0075945P.  
PR 20-MAR-1998; 98US-0078910P.  
PR 28-APR-1998; 98US-0083322P.  
PR 07-MAY-1998; 98US-0084600P.  
PR 28-MAY-1998; 98US-0087106P.  
PR 02-JUN-1998; 98US-0087607P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088021P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088026P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088030P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 05-JUN-1998; 98US-0088555P.  
PR 10-JUN-1998; 98US-0088734P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088742P.  
PR 10-JUN-1998; 98US-0088810P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088858P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089440P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089532P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089599P.  
PR 17-JUN-1998; 98US-0089600P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089801P.  
18-JUN-1998; 98US-0089907P.  
18-JUN-1998; 98US-0089908P.  
19-JUN-1998; 98US-0089947P.  
19-JUN-1998; 98US-0089948P.  
19-JUN-1998; 98US-0089952P.  
22-JUN-1998; 98US-0090246P.  
22-JUN-1998; 98US-0090252P.  
22-JUN-1998; 98US-0090254P.  
23-JUN-1998; 98US-0090349P.  
23-JUN-1998; 98US-0090355P.  
24-JUN-1998; 98US-0090429P.  
24-JUN-1998; 98US-0090431P.  
24-JUN-1998; 98US-0090435P.  
24-JUN-1998; 98US-0090444P.  
24-JUN-1998; 98US-0090445P.  
24-JUN-1998; 98US-0090472P.  
24-JUN-1998; 98US-0090535P.  
24-JUN-1998; 98US-0090540P.  
24-JUN-1998; 98US-0090542P.  
24-JUN-1998; 98US-0090577P.  
25-JUN-1998; 98US-0090676P.  
25-JUN-1998; 98US-0090678P.  
25-JUN-1998; 98US-0090690P.  
25-JUN-1998; 98US-0090694P.  
25-JUN-1998; 98US-0090695P.  
25-JUN-1998; 98US-0090696P.  
26-JUN-1998; 98US-0090862P.  
26-JUN-1998; 98US-0090863P.  
01-JUL-1998; 98US-0091360P.  
01-JUL-1998; 98US-0091544P.  
02-JUL-1998; 98US-0091478P.  
02-JUL-1998; 98US-0091519P.  
02-JUL-1998; 98US-0091626P.  
02-JUL-1998; 98US-0091628P.  
02-JUL-1998; 98US-0091633P.  
02-JUL-1998; 98US-0091646P.  
02-JUL-1998; 98US-0091673P.  
07-JUL-1998; 98US-0091978P.  
07-JUL-1998; 98US-0091982P.  
09-JUL-1998; 98US-0092182P.  
10-JUL-1998; 98US-0092472P.  
20-JUL-1998; 98US-0093339P.  
30-JUL-1998; 98US-0094651P.  
04-AUG-1998; 98US-0095282P.  
04-AUG-1998; 98US-0095285P.  
04-AUG-1998; 98US-0095301P.  
04-AUG-1998; 98US-0095302P.  
04-AUG-1998; 98US-0095318P.  
04-AUG-1998; 98US-0095321P.  
04-AUG-1998; 98US-0095325P.  
10-AUG-1998; 98US-0095916P.  
10-AUG-1998; 98US-0095929P.  
10-AUG-1998; 98US-0096012P.  
11-AUG-1998; 98US-0096143P.  
11-AUG-1998; 98US-0096146P.  
12-AUG-1998; 98US-0096329P.  
17-AUG-1998; 98US-0096757P.  
17-AUG-1998; 98US-0096766P.  
17-AUG-1998; 98US-0096768P.  
17-AUG-1998; 98US-0096773P.  
17-AUG-1998; 98US-0096791P.  
17-AUG-1998; 98US-0096867P.  
17-AUG-1998; 98US-0096891P.  
17-AUG-1998; 98US-0096894P.  
17-AUG-1998; 98US-0096895P.  
17-AUG-1998; 98US-0096897P.  
18-AUG-1998; 98US-0096949P.  
18-AUG-1998; 98US-0096950P.  
18-AUG-1998; 98US-0096959P.  
18-AUG-1998; 98US-0096960P.  
18-AUG-1998; 98US-0097022P.  
19-AUG-1998; 98US-0097141P.  
20-AUG-1998; 98US-0097218P.



PR 22-JUN-1998; 98US-0090252P.  
PR 23-JUN-1998; 98US-0090254P.  
PR 23-JUN-1998; 98US-0090349P.  
PR 23-JUN-1998; 98US-0090355P.  
PR 24-JUN-1998; 98US-0090423P.  
PR 24-JUN-1998; 98US-0090431P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090445P.  
PR 24-JUN-1998; 98US-0090472P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 24-JUN-1998; 98US-0090542P.  
PR 24-JUN-1998; 98US-0090557P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 01-JUL-1998; 98US-0091360P.  
PR 02-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091519P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091633P.  
PR 02-JUL-1998; 98US-0091646P.  
PR 02-JUL-1998; 98US-0091673P.  
PR 07-JUL-1998; 98US-0091978P.  
PR 07-JUL-1998; 98US-0091982P.  
PR 08-JUL-1998; 98US-0092182P.  
PR 10-JUL-1998; 98US-0092472P.  
PR 20-JUL-1998; 98US-0093339P.  
PR 30-JUL-1998; 98US-0094651P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 04-AUG-1998; 98US-0095285P.  
PR 04-AUG-1998; 98US-0095301P.  
PR 04-AUG-1998; 98US-0095302P.  
PR 04-AUG-1998; 98US-0095318P.  
PR 04-AUG-1998; 98US-0095321P.  
PR 04-AUG-1998; 98US-0095325P.  
PR 10-AUG-1998; 98US-0095916P.  
PR 10-AUG-1998; 98US-0095929P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 11-AUG-1998; 98US-0096143P.  
PR 11-AUG-1998; 98US-0096146P.  
PR 12-AUG-1998; 98US-0096329P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096768P.  
PR 17-AUG-1998; 98US-0096773P.  
PR 17-AUG-1998; 98US-0096791P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096894P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096950P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0096960P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 18-AUG-1998; 98US-0097141P.  
PR 19-AUG-1998; 98US-0097218P.  
PR 20-AUG-1998; 98US-0097661P.  
PR 24-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.

PR 26-AUG-1998; 98US-0097978P.  
PR 26-AUG-1998; 98US-0097979P.  
PR 26-AUG-1998; 98US-0097986P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 31-AUG-1998; 98US-0098525P.  
PR 16-SEP-1998; 98US-0100634P.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98US-0100858P.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 01-DEC-1998; 98WO-US025108.  
PR 22-DEC-1998; 98US-0113296P.  
PR 05-JAN-1999; 99WO-US000106.  
PR 08-MAR-1999; 99WO-US005028.  
PR 12-MAR-1999; 99US-0123957P.  
PR 02-JUN-1999; 99WO-US012252.  
PR 23-JUN-1999; 99US-0141037P.  
PR 07-JUL-1999; 99US-0143048P.  
PR 20-JUL-1999; 99US-0144758P.  
PR 26-JUL-1999; 99US-0145698P.  
PR 28-JUL-1999; 99US-0146222P.  
PR 17-AUG-1999; 99US-0149396P.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 08-OCT-1999; 99US-0158663P.  
PR 30-NOV-1999; 99WO-US028313.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030911.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 22-FEB-2000; 2000WO-US004114.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 15-MAY-2000; 2000WO-US013358.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015284.  
PR 23-JUN-2000; 2000US-0213637P.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US020331.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 07-SEP-2000; 2000US-0230978P.  
PR 08-NOV-2000; 2000WO-US030952.

Query Match 100.0%; Score 502; DB 6; Length 104;  
Best Local Similarity 100.0%; Pred. No. 1.4e-48;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKLAALLGLCVALLSCSSAAAFVLVGSKPVAAQVVALESAAAGAGTLANPLGTINPLKLL 60  
Db 1 MKLAALLGLCVALLSCSSAAAFVLVGSKPVAAQVVALESAAAGAGTLANPLGTINPLKLL 60  
Qy 61 LSSLGIPVNHLEIGSKQVVAELGPQAVGAVKALKALLGALTTFVG 104  
Db 61 LSSLGIPVNHLEIGSKQVVAELGPQAVGAVKALKALLGALTTFVG 104

RESULT 26  
ADA17936  
ID ADA17936 standard; protein; 104 AA.  
XX  
AC ADA17936;

XX 20-NOV-2003 (first entry)  
XX DT  
XX DE Human PRO1245 polypeptide.  
XX KW Human; PRO polypeptide; secreted protein; transmembrane protein;  
XX transgenic; tumour; cytostatic.  
XX OS Homo sapiens.  
XX FN US2003054987-A1.  
XX PD  
XX PF 20-MAR-2003.  
XX PP 14-NOV-2001; 2001US-00990443.  
PR 16-JUN-1997; 97US-0049787P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 05-NOV-1997; 97WO-US020069.  
PR 12-NOV-1997; 97US-0085186P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 24-NOV-1997; 97US-0066770P.  
PR 25-FEB-1998; 98US-0075945P.  
PR 20-MAR-1998; 98US-0078910P.  
PR 28-APR-1998; 98US-0083322P.  
PR 07-MAY-1998; 98US-0084600P.  
PR 28-MAY-1998; 98US-0087106P.  
PR 02-JUN-1998; 98US-0087607P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088021P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088026P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088030P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088734P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088742P.  
PR 10-JUN-1998; 98US-0088810P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088858P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 12-JUN-1998; 98US-0088765P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089440P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089532P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089599P.  
PR 17-JUN-1998; 98US-0089602P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089801P.  
PR 18-JUN-1998; 98US-0089907P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089947P.  
PR 19-JUN-1998; 98US-0089948P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 23-JUN-1998; 98US-0090349P.  
PR 23-JUN-1998; 98US-0090355P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090431P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090445P.  
PR 24-JUN-1998; 98US-0090472P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 24-JUN-1998; 98US-0090542P.  
PR 24-JUN-1998; 98US-0090557P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090894P.  
PR 25-JUN-1998; 98US-0090895P.  
PR 26-JUN-1998; 98US-0090896P.  
PR 26-JUN-1998; 98US-0090862P.  
PR 01-JUL-1998; 98US-0090863P.  
PR 01-JUL-1998; 98US-0091360P.  
PR 02-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091519P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091633P.  
PR 02-JUL-1998; 98US-0091646P.  
PR 02-JUL-1998; 98US-0091673P.  
PR 07-JUL-1998; 98US-0091978P.  
PR 07-JUL-1998; 98US-0091982P.  
PR 09-JUL-1998; 98US-0092182P.  
PR 10-JUL-1998; 98US-0092472P.  
PR 30-JUL-1998; 98US-0093339P.  
PR 04-AUG-1998; 98US-0094851P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 04-AUG-1998; 98US-0095285P.  
PR 04-AUG-1998; 98US-0095301P.  
PR 04-AUG-1998; 98US-0095302P.  
PR 04-AUG-1998; 98US-0095318P.  
PR 04-AUG-1998; 98US-0095321P.  
PR 10-AUG-1998; 98US-0095325P.  
PR 10-AUG-1998; 98US-0095916P.  
PR 10-AUG-1998; 98US-0095929P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 11-AUG-1998; 98US-0096143P.  
PR 12-AUG-1998; 98US-0096146P.  
PR 12-AUG-1998; 98US-0096329P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096773P.  
PR 17-AUG-1998; 98US-0096791P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096894P.  
PR 17-AUG-1998; 98US-0096895P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096950P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0096960P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 19-AUG-1998; 98US-0097141P.  
PR 20-AUG-1998; 98US-0097218P.  
PR 24-AUG-1998; 98US-0097661P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0097978P.  
PR 26-AUG-1998; 98US-0097979P.  
PR 26-AUG-1998; 98US-0097986P.

PR	26-AUG-1998;	98US-0098014P.	DE	Human secreted/transmembrane protein PRO1245.
PR	31-AUG-1998;	98US-0098525P.	XX	
PR	16-SEP-1998;	98US-0100634P.	KW	PRO; secreted protein; transmembrane protein;
PR	17-SEP-1998;	98US-0100634P.	KW	hypertrophy of neonatal heart; angiogenesis;
PR	16-SEP-1998;	98US-0100634P.	KW	vascular endothelial growth factor; VEGF-stimulated proliferation;
PR	17-SEP-1998;	98US-0100634P.	KW	endothelial cell; T-lymphocyte proliferation; retinal neuron;
PR	17-SEP-1998;	98US-0100634P.	KW	rod photoreceptor cell; c-fos induction; adipocyte cell;
PR	17-SEP-1998;	98US-0100634P.	KW	chondrocyte differentiation;
PR	01-OCT-1998;	98US-0100634P.	KW	pancreatic beta-cell precursor differentiation;
PR	01-OCT-1998;	98US-0100634P.	KW	cardiac insufficiency disorder; wound; cancerous tumour;
PR	05-DEC-1998;	98US-0100634P.	KW	retinal disorders; loss of sight; retinitis pigmentosa; kidney disorder;
PR	05-DEC-1998;	98US-0100634P.	KW	obesity; diabetes; hyperinsulinaemia; hypoparathyroidism; bone disorder;
PR	12-MAR-1999;	98US-0123957P.	KW	cartilage disorder; sports injury; arthritis; cancer; human.
PR	02-JUN-1999;	98US-0141037P.	XX	
PR	07-JUL-1999;	98US-0143048P.	XX	
PR	20-JUL-1999;	98US-0144758P.	OS	Homo sapiens.
PR	26-JUL-1999;	98US-0145698P.	XX	
PR	28-JUL-1999;	98US-0146222P.	PN	US2003054359-A1.
PR	17-AUG-1999;	98US-0149396P.	XX	
PR	15-SEP-1999;	98US-0149396P.	XX	
PR	15-SEP-1999;	98US-0149396P.	XX	
PR	08-OCT-1999;	98US-0158663P.	XX	
PR	30-NOV-1999;	98US-0202831P.	XX	
PR	01-DEC-1999;	98US-0202831P.	XX	
PR	01-DEC-1999;	98US-0202831P.	XX	
PR	16-DEC-1999;	98US-0203009P.	XX	
PR	20-DEC-1999;	98US-0203091P.	XX	
PR	05-JAN-2000;	2000US-0200219P.	XX	
PR	06-JAN-2000;	2000US-0200376P.	XX	
PR	11-FEB-2000;	2000US-0200356P.	XX	
PR	22-FEB-2000;	2000US-0200441P.	XX	
PR	24-FEB-2000;	2000US-0200491P.	XX	
PR	24-FEB-2000;	2000US-0200500P.	XX	
PR	02-MAR-2000;	2000US-0200584P.	XX	
PR	18-MAR-2000;	2000US-0200631P.	XX	
PR	18-MAR-2000;	2000US-0200688P.	XX	
PR	30-MAR-2000;	2000US-0200737P.	XX	
PR	30-MAR-2000;	2000US-0200843P.	XX	
PR	15-MAY-2000;	2000US-0201335P.	XX	
PR	17-MAY-2000;	2000US-0201370P.	XX	
PR	22-MAY-2000;	2000US-0201404P.	XX	
PR	30-MAY-2000;	2000US-0201491P.	XX	
PR	02-JUN-2000;	2000US-0201526P.	XX	
PR	23-JUN-2000;	2000US-0201637P.	XX	
PR	28-JUL-2000;	2000US-0202071P.	XX	
PR	11-AUG-2000;	2000US-0202203P.	XX	
PR	23-AUG-2000;	2000US-0202352P.	XX	
PR	24-AUG-2000;	2000US-0202352P.	XX	
PR	07-SEP-2000;	2000US-0203078P.	XX	
PR	08-NOV-2000;	2000US-0203095P.	XX	
PR	26-AUG-1998;	98US-0098014P.	DE	Human secreted/transmembrane protein PRO1245.
PR	31-AUG-1998;	98US-0098525P.	XX	
PR	16-SEP-1998;	98US-0100634P.	KW	PRO; secreted protein; transmembrane protein;
PR	17-SEP-1998;	98US-0100634P.	KW	hypertrophy of neonatal heart; angiogenesis;
PR	16-SEP-1998;	98US-0100634P.	KW	vascular endothelial growth factor; VEGF-stimulated proliferation;
PR	17-SEP-1998;	98US-0100634P.	KW	endothelial cell; T-lymphocyte proliferation; retinal neuron;
PR	17-SEP-1998;	98US-0100634P.	KW	rod photoreceptor cell; c-fos induction; adipocyte cell;
PR	17-SEP-1998;	98US-0100634P.	KW	chondrocyte differentiation;
PR	01-OCT-1998;	98US-0100634P.	KW	pancreatic beta-cell precursor differentiation;
PR	01-OCT-1998;	98US-0100634P.	KW	cardiac insufficiency disorder; wound; cancerous tumour;
PR	05-DEC-1998;	98US-0100634P.	KW	retinal disorders; loss of sight; retinitis pigmentosa; kidney disorder;
PR	05-DEC-1998;	98US-0100634P.	KW	obesity; diabetes; hyperinsulinaemia; hypoparathyroidism; bone disorder;
PR	12-MAR-1999;	98US-0123957P.	KW	cartilage disorder; sports injury; arthritis; cancer; human.
PR	02-JUN-1999;	98US-0141037P.	XX	
PR	07-JUL-1999;	98US-0143048P.	XX	
PR	20-JUL-1999;	98US-0144758P.	OS	Homo sapiens.
PR	26-JUL-1999;	98US-0145698P.	XX	
PR	28-JUL-1999;	98US-0146222P.	PN	US2003054359-A1.
PR	17-AUG-1999;	98US-0149396P.	XX	
PR	15-SEP-1999;	98US-0149396P.	XX	
PR	15-SEP-1999;	98US-0149396P.	XX	
PR	08-OCT-1999;	98US-0158663P.	XX	
PR	30-NOV-1999;	98US-0202831P.	XX	
PR	01-DEC-1999;	98US-0202831P.	XX	
PR	01-DEC-1999;	98US-0202831P.	XX	
PR	16-DEC-1999;	98US-0203009P.	XX	
PR	20-DEC-1999;	98US-0203091P.	XX	
PR	05-JAN-2000;	2000US-0200219P.	XX	
PR	06-JAN-2000;	2000US-0200376P.	XX	
PR	11-FEB-2000;	2000US-0200356P.	XX	
PR	22-FEB-2000;	2000US-0200441P.	XX	
PR	24-FEB-2000;	2000US-0200491P.	XX	
PR	24-FEB-2000;	2000US-0200500P.	XX	
PR	02-MAR-2000;	2000US-0200584P.	XX	
PR	18-MAR-2000;	2000US-0200631P.	XX	
PR	18-MAR-2000;	2000US-0200688P.	XX	
PR	30-MAR-2000;	2000US-0200737P.	XX	
PR	30-MAR-2000;	2000US-0200843P.	XX	
PR	15-MAY-2000;	2000US-0201335P.	XX	
PR	17-MAY-2000;	2000US-0201370P.	XX	
PR	22-MAY-2000;	2000US-0201404P.	XX	
PR	30-MAY-2000;	2000US-0201491P.	XX	
PR	02-JUN-2000;	2000US-0201526P.	XX	
PR	23-JUN-2000;	2000US-0201637P.	XX	
PR	28-JUL-2000;	2000US-0202071P.	XX	
PR	11-AUG-2000;	2000US-0202203P.	XX	
PR	23-AUG-2000;	2000US-0202352P.	XX	
PR	24-AUG-2000;	2000US-0202352P.	XX	
PR	07-SEP-2000;	2000US-0203078P.	XX	
PR	08-NOV-2000;	2000US-0203095P.	XX	
Qy	1 MKLALGLCVLSCSAARFLVGSAPVAPVVALESAAEAGAGTLANPLGLNPLKLL 60			
Db	1 MKLALGLCVLSCSAARFLVGSAPVAPVVALESAAEAGAGTLANPLGLNPLKLL 60			
Qy	61 LSSLGIPVNHLEIGSKCVKVAELQPOAVGAVKALKALLGALTTFG 104			
Db	61 LSSLGIPVNHLEIGSKCVKVAELQPOAVGAVKALKALLGALTTFG 104			
RESULT 27				
ADA28044				
ID ADA28044	standard; protein; 104 AA.			
XX AC				
XX ADA28044;				
XX				
DT 20-NOV-2003	(first entry)			
XX				





[illegible]

PR	26-AUG-1998;	98US-0098014P.	XX	US2003059780-A1.
PR	31-AUG-1998;	98US-0098525P.	PN	
PR	16-SEP-1998;	98US-0100634P.	XX	27-MAR-2003.
PR	16-SEP-1998;	98WO-US019330.	PD	
PR	17-SEP-1998;	98US-0100858P.	XX	
PR	17-SEP-1998;	98WO-US019437.	PF	14-NOV-2001; 2001US-00991854.
PR	07-OCT-1998;	98WO-US021141.	XX	
PR	01-DEC-1998;	98WO-US025108.	PR	16-JUN-1997; 97US-0049787P.
PR	22-DEC-1998;	98US-0113296P.	PR	17-OCT-1997; 97US-0062250P.
PR	05-JAN-1999;	99WO-US000106.	PR	05-NOV-1997; 97WO-US020069.
PR	08-MAR-1999;	99WO-US005028.	PR	12-NOV-1997; 97US-0065186P.
PR	12-MAR-1999;	99US-0123957P.	PR	13-NOV-1997; 97US-0065311P.
PR	02-JUN-1999;	99WO-US012252.	PR	24-NOV-1997; 97US-0066770P.
PR	23-JUN-1999;	99US-0141037P.	PR	25-FEB-1998; 98US-0075945P.
PR	07-JUL-1999;	99US-0143048P.	PR	25-FEB-1998; 98US-0078910P.
PR	20-JUL-1999;	99US-0144758P.	PR	28-MAR-1998; 98US-0083322P.
PR	26-JUL-1999;	99US-0145698P.	PR	28-APR-1998; 98US-0083322P.
PR	28-JUL-1999;	99US-0146222P.	PR	07-MAY-1998; 98US-0084600P.
PR	17-AUG-1999;	99US-0149396P.	PR	28-MAY-1998; 98US-0087106P.
PR	15-SEP-1999;	99WO-US021090.	PR	02-JUN-1998; 98US-0087607P.
PR	15-SEP-1999;	99WO-US021547.	PR	02-JUN-1998; 98US-0087609P.
PR	08-OCT-1999;	99US-0158663P.	PR	02-JUN-1998; 98US-0087759P.
PR	30-NOV-1999;	99WO-US028313.	PR	03-JUN-1998; 98US-0087827P.
PR	01-DEC-1999;	99WO-US028301.	PR	04-JUN-1998; 98US-0088021P.
PR	16-DEC-1999;	99WO-US028634.	PR	04-JUN-1998; 98US-0088025P.
PR	20-DEC-1999;	99WO-US030095.	PR	04-JUN-1998; 98US-0088026P.
PR	05-JAN-2000;	2000WO-US000219.	PR	04-JUN-1998; 98US-0088030P.
PR	06-JAN-2000;	2000WO-US000376.	PR	04-JUN-1998; 98US-0088033P.
PR	11-FEB-2000;	2000WO-US003565.	PR	04-JUN-1998; 98US-0088326P.
PR	18-FEB-2000;	2000WO-US004341.	PR	05-JUN-1998; 98US-0088202P.
PR	22-FEB-2000;	2000WO-US004414.	PR	05-JUN-1998; 98US-0088212P.
PR	24-FEB-2000;	2000WO-US004914.	PR	05-JUN-1998; 98US-0088217P.
PR	02-MAR-2000;	2000WO-US005004.	PR	09-JUN-1998; 98US-0088655P.
PR	02-MAR-2000;	2000WO-US005841.	PR	10-JUN-1998; 98US-0088734P.
PR	10-MAR-2000;	2000WO-US006319.	PR	10-JUN-1998; 98US-0088738P.
PR	15-MAR-2000;	2000WO-US006884.	PR	10-JUN-1998; 98US-0088742P.
PR	20-MAR-2000;	2000WO-US007377.	PR	10-JUN-1998; 98US-0088810P.
PR	30-MAR-2000;	2000WO-US008439.	PR	10-JUN-1998; 98US-0088824P.
PR	15-MAY-2000;	2000WO-US013358.	PR	10-JUN-1998; 98US-0088826P.
PR	17-MAY-2000;	2000WO-US013705.	PR	11-JUN-1998; 98US-0088858P.
PR	22-MAY-2000;	2000WO-US014042.	PR	11-JUN-1998; 98US-0088861P.
PR	30-MAY-2000;	2000WO-US014941.	PR	11-JUN-1998; 98US-0088876P.
PR	02-JUN-2000;	2000WO-US015264.	PR	12-JUN-1998; 98US-0089105P.
PR	23-JUN-2000;	2000US-0213637P.	PR	16-JUN-1998; 98US-0089440P.
Query Match 100.0%; Score 502; DB 6; Length 104;				
Best Local Similarity 100.0%; Pred. No. 1.4e-48;				
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	MKLAALLGLCVALLSCSSAAAFVLSAKPVAQPVAALESAAEAGAGTLANPLGTINPLKLL 60	PR	16-JUN-1998; 98US-0089512P.
Db	1	MKLAALLGLCVALLSCSSAAAFVLSAKPVAQPVAALESAAEAGAGTLANPLGTINPLKLL 60	PR	16-JUN-1998; 98US-0089514P.
Qy	61	LSSLGIPVNHILGSKCVAELGPQAVGAVKALKALLGALTVEG 104	PR	17-JUN-1998; 98US-0089532P.
Db	61	LSSLGIPVNHILGSKCVAELGPQAVGAVKALKALLGALTVEG 104	PR	17-JUN-1998; 98US-0089538P.
RESULT 29				
ADA38849				
ID	ADA38849 standard; protein; 104 AA.			
XX	ADA38849;			
XX	20-NOV-2003. (first entry)			
DT	Human secreted/transmembrane protein PRO1245.			
DE	PRO; secreted protein; transmembrane protein; gene therapy; tumour;			
XX	cancer; human; colon cancer; lung cancer; breast cancer.			
KW	Homo sapiens.			
XX				
OS				

PR 24-JUN-1998; 98US-0090542P.  
PR 24-JUN-1998; 98US-0090557P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 26-JUN-1998; 98US-0090862P.  
PR 01-JUL-1998; 98US-0091360P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091519P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091633P.  
PR 02-JUL-1998; 98US-0091646P.  
PR 02-JUL-1998; 98US-0091671P.  
PR 07-JUL-1998; 98US-0091978P.  
PR 07-JUL-1998; 98US-0091982P.  
PR 09-JUL-1998; 98US-0092182P.  
PR 10-JUL-1998; 98US-0092472P.  
PR 20-JUL-1998; 98US-0093339P.  
PR 20-JUL-1998; 98US-0094651P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 04-AUG-1998; 98US-0095288P.  
PR 04-AUG-1998; 98US-0095301P.  
PR 04-AUG-1998; 98US-0095302P.  
PR 04-AUG-1998; 98US-0095318P.  
PR 04-AUG-1998; 98US-0095321P.  
PR 04-AUG-1998; 98US-0095325P.  
PR 10-AUG-1998; 98US-0095916P.  
PR 10-AUG-1998; 98US-0095922P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 11-AUG-1998; 98US-0096143P.  
PR 11-AUG-1998; 98US-0096146P.  
PR 12-AUG-1998; 98US-0096329P.  
PR 12-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096768P.  
PR 17-AUG-1998; 98US-0096773P.  
PR 17-AUG-1998; 98US-0096791P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096894P.  
PR 17-AUG-1998; 98US-0096895P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096849P.  
PR 18-AUG-1998; 98US-0096950P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0096960P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 19-AUG-1998; 98US-0097141P.  
PR 20-AUG-1998; 98US-0097218P.  
PR 20-AUG-1998; 98US-0097661P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0097978P.  
PR 26-AUG-1998; 98US-0097979P.  
PR 26-AUG-1998; 98US-0097986P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 31-AUG-1998; 98US-0098525P.  
PR 16-SEP-1998; 98US-0100634P.  
PR 16-SEP-1998; 98US-0100630P.  
PR 17-SEP-1998; 98US-0100858P.  
PR 17-SEP-1998; 98US-0100858P.  
PR 07-OCT-1998; 98US-0100858P.  
PR 01-DEC-1998; 98US-0100858P.  
PR 22-DEC-1998; 98US-0113296P.

PR 05-JAN-1999; 99WO-US000106.  
PR 08-MAR-1999; 99WO-US0005028.  
PR 12-MAR-1999; 99WO-0123957P.  
PR 02-JUN-1999; 99WO-US012252.  
PR 23-JUN-1999; 99US-0141037P.  
PR 07-JUL-1999; 99US-0143048P.  
PR 20-JUL-1999; 99US-0144758P.  
PR 26-JUL-1999; 99US-0145698P.  
PR 28-JUL-1999; 99US-0146222P.  
PR 17-AUG-1999; 99US-0146396P.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 08-OCT-1999; 99US-0158663P.  
PR 30-NOV-1999; 99WO-US028313.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030511.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US004341.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 15-MAY-2000; 2000WO-US013358.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000US-0213637P.  
PR 23-JUN-2000; 2000US-0213637P.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023528.  
PR 07-SEP-2000; 2000US-0230978P.  
PR 08-NOV-2000; 2000WO-US030952.

Query Match 100.0%; Score 502; DB 6; Length 104;  
Best Local Similarity 100.0%; Pred. No. 1.4e-48;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLAALLGLCVALSAAFLVGSAPVQVVALESAAEAGAGTLANPLGTINPLKLL 60  
Db 1 MKLAALLGLCVALSAAFLVGSAPVQVVALESAAEAGAGTLANPLGTINPLKLL 60

QY 61 LSSLGIPVNHLEGSQKCVAEIGPOAVGAVKALKALIGALTVFG 104  
Db 61 LSSLGIPVNHLEGSQKCVAEIGPOAVGAVKALKALIGALTVFG 104

RESULT 30  
ADAS2970

ID ADAS2970 standard; protein; 104 AA.

XX ADAS2970;

XX 20-NOV-2003 (first entry)

XX Human secreted/transmembrane protein PRO1245.

XX PRO; secreted protein; transmembrane protein;  
XX hypertrophy of neonatal heart; angiogenesis;  
XX vascular endothelial growth factor; VEGF-stimulated proliferation;  
XX endothelial cell; T-lymphocyte proliferation; retinal neuron;  
XX c-fos induction; adipocyte cell; chondrocyte differentiation;  
XX pancreatic beta-cell precursor differentiation; gene therapy; tumour;  
XX cancer; human; colon cancer; lung cancer; breast cancer;

KW rod photoreceptor cell.  
XX Homo sapiens.  
OS US2003060407-A1.  
XX 27-MAR-2003.  
XX 14-NOV-2001; 2001US-00990440.  
XX 16-JUN-1997; 97US-0649787P.  
PR 17-OCT-1997; 97US-062250P.  
PR 05-NOV-1997; 97WO-US020069.  
PR 12-NOV-1997; 97US-065186P.  
PR 13-NOV-1997; 97US-065311P.  
PR 24-NOV-1997; 97US-066770P.  
PR 25-FEB-1998; 98US-0075945P.  
PR 20-MAR-1998; 98US-0078910P.  
PR 28-APR-1998; 98US-0083322P.  
PR 07-MAY-1998; 98US-0084600P.  
PR 28-MAY-1998; 98US-0087106P.  
PR 02-JUN-1998; 98US-0087607P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 02-JUN-1998; 98US-0087827P.  
PR 03-JUN-1998; 98US-0088021P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088026P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088030P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088467P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088734P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088742P.  
PR 10-JUN-1998; 98US-0088810P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088858P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089440P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089532P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089599P.  
PR 17-JUN-1998; 98US-0089600P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089801P.  
PR 18-JUN-1998; 98US-0089907P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089947P.  
PR 19-JUN-1998; 98US-0089948P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 23-JUN-1998; 98US-0090349P.  
PR 23-JUN-1998; 98US-0090355P.  
PR 24-JUN-1998; 98US-0090423P.  
PR 24-JUN-1998; 98US-0090431P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090445P.  
PR 24-JUN-1998; 98US-0090472P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 24-JUN-1998; 98US-0090542P.  
PR 24-JUN-1998; 98US-0090557P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 01-JUL-1998; 98US-0091360P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091519P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091633P.  
PR 02-JUL-1998; 98US-0091646P.  
PR 02-JUL-1998; 98US-0091673P.  
PR 07-JUL-1998; 98US-0091978P.  
PR 07-JUL-1998; 98US-0091982P.  
PR 09-JUL-1998; 98US-0092182P.  
PR 10-JUL-1998; 98US-0092472P.  
PR 20-JUL-1998; 98US-0093339P.  
PR 30-JUL-1998; 98US-0094651P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 04-AUG-1998; 98US-0095285P.  
PR 04-AUG-1998; 98US-0095301P.  
PR 04-AUG-1998; 98US-0095302P.  
PR 04-AUG-1998; 98US-0095318P.  
PR 04-AUG-1998; 98US-0095321P.  
PR 04-AUG-1998; 98US-0095325P.  
PR 10-AUG-1998; 98US-0095916P.  
PR 10-AUG-1998; 98US-0095929P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 11-AUG-1998; 98US-0096143P.  
PR 11-AUG-1998; 98US-0096146P.  
PR 12-AUG-1998; 98US-0096329P.  
PR 12-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096768P.  
PR 17-AUG-1998; 98US-0096773P.  
PR 17-AUG-1998; 98US-0096791P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096894P.  
PR 17-AUG-1998; 98US-0096895P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096950P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0096960P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 18-AUG-1998; 98US-0097141P.  
PR 19-AUG-1998; 98US-0097218P.  
PR 20-AUG-1998; 98US-0097261P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0097978P.  
PR 26-AUG-1998; 98US-0097979P.  
PR 26-AUG-1998; 98US-0097986P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 31-AUG-1998; 98US-0098525P.  
PR 16-SEP-1998; 98US-0100634P.  
PR 17-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98US-0100858P.  
PR 17-SEP-1998; 98WO-US019437.

PR	07-OCT-1998;	98WO-US021141.	PN	US2003044806-A1.
PR	01-DEC-1998;	98WO-US025108.	XX	
PR	22-DEC-1998;	98US-0113298P.	PD	
PR	05-JAN-1999;	99WO-US000106.	XX	06-MAR-2003.
PR	08-MAR-1999;	99WO-US005028.	PP	15-NOV-2001; 2001US-00998156.
PR	12-MAR-1999;	99WO-US012357P.	XX	
PR	02-JUN-1999;	99WO-US012252.	PR	16-JUN-1997; 97US-0049787P.
PR	23-JUN-1999;	99US-0141037P.	PR	17-OCT-1997; 97US-0062250P.
PR	07-JUL-1999;	99US-0143048P.	PR	05-NOV-1997; 97WO-US020069.
PR	20-JUL-1999;	99US-0144758P.	PR	12-NOV-1997; 97US-0065186P.
PR	26-JUL-1999;	99US-0145698P.	PR	13-NOV-1997; 97US-0065311P.
PR	28-JUL-1999;	99US-0146222P.	PR	24-NOV-1997; 97US-0066770P.
PR	17-AUG-1999;	99US-0149396P.	PR	25-FEB-1998; 98US-0075945P.
PR	15-SEP-1999;	99WO-US021090.	PR	20-MAR-1998; 98US-0078910P.
PR	08-OCT-1999;	99WO-US021547.	PR	28-APR-1998; 98US-0083342P.
PR	15-SEP-1999;	99US-0158663P.	PR	07-MAY-1998; 98US-0084600P.
PR	30-NOV-1999;	99WO-US028313.	PR	28-MAY-1998; 98US-0087106P.
PR	01-DEC-1999;	99WO-US028301.	PR	02-JUN-1998; 98US-0087607P.
PR	01-DEC-1999;	99WO-US028634.	PR	02-JUN-1998; 98US-0087609P.
PR	16-DEC-1999;	99WO-US030095.	PR	02-JUN-1998; 98US-0087759P.
PR	20-DEC-1999;	99WO-US030911.	PR	03-JUN-1998; 98US-0087827P.
PR	05-JAN-2000;	2000WO-US000219.	PR	04-JUN-1998; 98US-0088012P.
PR	06-JAN-2000;	2000WO-US000376.	PR	04-JUN-1998; 98US-0088026P.
PR	11-FEB-2000;	2000WO-US003565.	PR	04-JUN-1998; 98US-0088038P.
PR	18-FEB-2000;	2000WO-US004341.	PR	04-JUN-1998; 98US-0088039P.
PR	22-FEB-2000;	2000WO-US004414.	PR	04-JUN-1998; 98US-0088030P.
PR	24-FEB-2000;	2000WO-US004914.	PR	04-JUN-1998; 98US-0088033P.
PR	24-FEB-2000;	2000WO-US005004.	PR	04-JUN-1998; 98US-0088326P.
PR	02-MAR-2000;	2000WO-US005841.	PR	05-JUN-1998; 98US-0088167P.
PR	10-MAR-2000;	2000WO-US006319.	PR	05-JUN-1998; 98US-0088202P.
PR	15-MAR-2000;	2000WO-US006884.	PR	05-JUN-1998; 98US-0088212P.
PR	20-MAR-2000;	2000WO-US007377.	PR	05-JUN-1998; 98US-0088217P.
PR	30-MAR-2000;	2000WO-US008439.	PR	09-JUN-1998; 98US-0088655P.
PR	15-MAY-2000;	2000WO-US013358.	PR	10-JUN-1998; 98US-0088734P.
PR	17-MAY-2000;	2000WO-US013705.	PR	10-JUN-1998; 98US-0088738P.
PR	22-MAY-2000;	2000WO-US014042.	PR	10-JUN-1998; 98US-0088810P.
PR	30-MAY-2000;	2000WO-US014941.	PR	10-JUN-1998; 98US-0088824P.
PR	02-JUN-2000;	2000WO-US015264.	PR	10-JUN-1998; 98US-0088826P.
PR	23-JUN-2000;	2000US-0213637P.	PR	11-JUN-1998; 98US-0088858P.
Query Match 100.0%; Score 502; DB 6; Length 104;			PR	11-JUN-1998; 98US-0088861P.
Best Local Similarity 100.0%; Pred. No. 1.4e-48;			PR	11-JUN-1998; 98US-0088867P.
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			PR	12-JUN-1998; 98US-0089105P.
Qy	1	MKLAALLGLCVALS	SSAAFLVGS	AKPVAQPVAALESAAEAGAGTLANPLGTINPLKLL 60
Db	1	MKLAALLGLCVALS	SSAAFLVGS	AKPVAQPVAALESAAEAGAGTLANPLGTINPLKLL 60
Qy	61	LSSLGIPVNHLEIGSKCV	AEILGPQAVGAVKALKALLGALT	VFVG 104
Db	61	LSSLGIPVNHLEIGSKCV	AEILGPQAVGAVKALKALLGALT	VFVG 104
RESULT 31				
ABO53232				
ID ABO53232 standard; protein; 104 AA.				
XX	AC	ABO53232;		
XX	AC	ABO53232;		
XX	DT	14-OCT-2003 (first entry)		
XX	DE	Human secreted/transmembrane protein PRO1245.		
KW	KW	Human; secreted protein; transmembrane protein; PRO;		
KW	KW	adrenal cortical capillary endothelial cell; angiogenesis; wound healing;		
KW	KW	diabetes; obesity; hyper-insulinaemia; hypo-insulinaemia;		
KW	KW	chondrocyte redifferentiation; bone disorder; cartilage disorder;		
KW	KW	sports injury; arthritis; kidney mesangial cell proliferation;		
KW	KW	kidney disorder; Berger disease; neuropathy; coeliac disease;		
XX	OS	Dermatitis herpetiformis; Crohn's disease; tumour; cancer.		
XX	OS	Homo sapiens.		

PR 24-JUN-1998; 98US-0090557P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 25-JUN-1998; 98US-0090697P.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 01-JUL-1998; 98US-0091360P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091513P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091633P.  
PR 02-JUL-1998; 98US-0091646P.  
PR 02-JUL-1998; 98US-0091673P.  
PR 07-JUL-1998; 98US-0091978P.  
PR 07-JUL-1998; 98US-0091982P.  
PR 09-JUL-1998; 98US-0092182P.  
PR 10-JUL-1998; 98US-0092472P.  
PR 20-JUL-1998; 98US-0093339P.  
PR 30-JUL-1998; 98US-0094651P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 04-AUG-1998; 98US-0095285P.  
PR 04-AUG-1998; 98US-0095301P.  
PR 04-AUG-1998; 98US-0095302P.  
PR 04-AUG-1998; 98US-0095318P.  
PR 04-AUG-1998; 98US-0095321P.  
PR 04-AUG-1998; 98US-0095325P.  
PR 10-AUG-1998; 98US-0095916P.  
PR 10-AUG-1998; 98US-0095929P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 11-AUG-1998; 98US-0096146P.  
PR 12-AUG-1998; 98US-0096329P.  
PR 13-AUG-1998; 98US-0096413P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096768P.  
PR 17-AUG-1998; 98US-0096773P.  
PR 17-AUG-1998; 98US-0096791P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096894P.  
PR 17-AUG-1998; 98US-0096895P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096950P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0096960P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 19-AUG-1998; 98US-0097141P.  
PR 20-AUG-1998; 98US-0097218P.  
PR 24-AUG-1998; 98US-0097661P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0097978P.  
PR 26-AUG-1998; 98US-0097979P.  
PR 26-AUG-1998; 98US-0097986P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 31-AUG-1998; 98US-0098525P.  
PR 16-SEP-1998; 98US-0100634P.  
PR 16-SEP-1998; 98US-0100634P.  
PR 17-SEP-1998; 98US-0100858P.  
PR 17-SEP-1998; 98US-0100858P.  
PR 07-OCT-1998; 98US-0100858P.  
PR 01-DEC-1998; 98US-0100858P.  
PR 01-DEC-1998; 98US-0100858P.  
PR 02-DEC-1998; 98US-0100858P.  
PR 05-JAN-1999; 99US-0098972P.

PR 08-MAR-1999; 99WO-US005028.  
PR 12-MAR-1999; 99US-0123957P.  
PR 02-JUN-1999; 99WO-US012252.  
PR 23-JUN-1999; 99US-0141037P.  
PR 07-JUL-1999; 99US-0143048P.  
PR 20-JUL-1999; 99US-0144758P.  
PR 26-JUL-1999; 99US-0145698P.  
PR 28-JUL-1999; 99US-0146222P.  
PR 17-AUG-1999; 99US-0149396P.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 08-OCT-1999; 99US-0158663P.  
PR 30-NOV-1999; 99WO-US028313.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030911.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 15-MAY-2000; 2000WO-US013358.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 23-JUN-2000; 2000US-0213637P.  
PR 28-JUL-2000; 2000WO-US020710.

Query Match 100.0%; Score 502; DB 7; Length 104;  
Best Local Similarity 100.0%; Pred. No. 1.4e-48;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKLAALGLCVLSCSSAAFLVGSAPVQVVALESAAEAGAGTLANPLGTLNPKLL 60  
Db 1 MKLAALGLCVLSCSSAAFLVGSAPVQVVALESAAEAGAGTLANPLGTLNPKLL 60  
Qy 61 LSSLGIPVNHLEGSQKVAELGPQAVGAVKALKALLGALTVEG 104  
Db 61 LSSLGIPVNHLEGSQKVAELGPQAVGAVKALKALLGALTVEG 104

## RESULT 32

ADA22531  
ID ADA22531 standard; protein; 104 AA.

XX ADA22531;

XX 20-NOV-2003 (first entry)

XX Human secreted/transmembrane polypeptide PRO1245.

XX human; tumour; cancer; colorectal cancer; gene therapy;  
KW chondrocyte differentiation; VEGF inhibition;  
KW vascular endothelial growth factor; Alzheimer's disease;  
KW Parkinson's disease; atherosclerosis; cystic fibrosis;  
XX multiple sclerosis; ovarian cancer; tissue typing.

OS Homo sapiens.

XX US2003040473-A1.

XX 27-FEB-2003.

XX 19-NOV-2001; 2001US-00989726.

XX	16-JUN-1997;	97US-0049787P.	98US-0090695P.	PR	25-JUN-1998;	98US-0090695P.
PR	17-OCT-1997;	97US-0062250P.	98US-0090696P.	PR	25-JUN-1998;	98US-0090696P.
PR	05-NOV-1997;	97WO-US020069.	98US-0090882P.	PR	26-JUN-1998;	98US-0090882P.
PR	12-NOV-1997;	97US-0065186P.	98US-0090863P.	PR	26-JUN-1998;	98US-0090863P.
PR	11-NOV-1997;	97US-0065311P.	98US-0091360P.	PR	01-JUL-1998;	98US-0091360P.
PR	24-NOV-1997;	97US-0065770P.	98US-0091544P.	PR	01-JUL-1998;	98US-0091544P.
PR	25-FEB-1998;	98US-0075945P.	98US-0091478P.	PR	02-JUL-1998;	98US-0091478P.
PR	20-MAR-1998;	98US-0076910P.	98US-0091519P.	PR	02-JUL-1998;	98US-0091519P.
PR	28-APR-1998;	98US-0083322P.	98US-0091626P.	PR	02-JUL-1998;	98US-0091626P.
PR	07-MAY-1998;	98US-0084600P.	98US-0091628P.	PR	02-JUL-1998;	98US-0091628P.
PR	28-MAY-1998;	98US-0087106P.	98US-0091633P.	PR	02-JUL-1998;	98US-0091633P.
PR	02-JUN-1998;	98US-0087607P.	98US-0091646P.	PR	02-JUL-1998;	98US-0091646P.
PR	02-JUN-1998;	98US-0087759P.	98US-0091673P.	PR	02-JUL-1998;	98US-0091673P.
PR	03-JUN-1998;	98US-0087827P.	98US-0091982P.	PR	07-JUL-1998;	98US-0091982P.
PR	04-JUN-1998;	98US-0088021P.	98US-0091822P.	PR	07-JUL-1998;	98US-0091822P.
PR	04-JUN-1998;	98US-0088025P.	98US-0092472P.	PR	09-JUL-1998;	98US-0092472P.
PR	04-JUN-1998;	98US-0088026P.	98US-0093339P.	PR	20-JUL-1998;	98US-0093339P.
PR	04-JUN-1998;	98US-0088028P.	98US-0094651P.	PR	30-JUL-1998;	98US-0094651P.
PR	04-JUN-1998;	98US-0088030P.	98US-0095282P.	PR	04-AUG-1998;	98US-0095282P.
PR	04-JUN-1998;	98US-0088033P.	98US-0095285P.	PR	04-AUG-1998;	98US-0095285P.
PR	04-JUN-1998;	98US-0088326P.	98US-0095301P.	PR	04-AUG-1998;	98US-0095301P.
PR	05-JUN-1998;	98US-0088167P.	98US-0095302P.	PR	04-AUG-1998;	98US-0095302P.
PR	05-JUN-1998;	98US-0088202P.	98US-0095318P.	PR	04-AUG-1998;	98US-0095318P.
PR	05-JUN-1998;	98US-0088217P.	98US-0095321P.	PR	04-AUG-1998;	98US-0095321P.
PR	09-JUN-1998;	98US-0088655P.	98US-0095325P.	PR	10-AUG-1998;	98US-0095325P.
PR	10-JUN-1998;	98US-0088734P.	98US-0095916P.	PR	10-AUG-1998;	98US-0095916P.
PR	10-JUN-1998;	98US-0088738P.	98US-0095922P.	PR	10-AUG-1998;	98US-0095922P.
PR	10-JUN-1998;	98US-0088742P.	98US-0096012P.	PR	10-AUG-1998;	98US-0096012P.
PR	10-JUN-1998;	98US-0088810P.	98US-0096143P.	PR	11-AUG-1998;	98US-0096143P.
PR	10-JUN-1998;	98US-0088824P.	98US-0096146P.	PR	11-AUG-1998;	98US-0096146P.
PR	11-JUN-1998;	98US-0088858P.	98US-0096329P.	PR	12-AUG-1998;	98US-0096329P.
PR	11-JUN-1998;	98US-0088861P.	98US-0096757P.	PR	17-AUG-1998;	98US-0096757P.
PR	12-JUN-1998;	98US-0088766P.	98US-0096766P.	PR	17-AUG-1998;	98US-0096766P.
PR	12-JUN-1998;	98US-0089105P.	98US-0096768P.	PR	17-AUG-1998;	98US-0096768P.
PR	16-JUN-1998;	98US-0089440P.	98US-0096773P.	PR	17-AUG-1998;	98US-0096773P.
PR	16-JUN-1998;	98US-0089512P.	98US-0096791P.	PR	17-AUG-1998;	98US-0096791P.
PR	16-JUN-1998;	98US-0089514P.	98US-0096867P.	PR	17-AUG-1998;	98US-0096867P.
PR	17-JUN-1998;	98US-0089532P.	98US-0096891P.	PR	17-AUG-1998;	98US-0096891P.
PR	17-JUN-1998;	98US-0089538P.	98US-0096894P.	PR	17-AUG-1998;	98US-0096894P.
PR	17-JUN-1998;	98US-0089598P.	98US-0096895P.	PR	17-AUG-1998;	98US-0096895P.
PR	17-JUN-1998;	98US-0089599P.	98US-0096897P.	PR	17-AUG-1998;	98US-0096897P.
PR	17-JUN-1998;	98US-0089600P.	98US-0096949P.	PR	18-AUG-1998;	98US-0096949P.
PR	17-JUN-1998;	98US-0089653P.	98US-0096950P.	PR	18-AUG-1998;	98US-0096950P.
PR</						



PR	20-JUL-1999;	99US-0144758P.	PR	05-NOV-1997;	97WO-US020069.
PR	26-JUL-1999;	99US-0145598P.	PR	12-NOV-1997;	97US-0065186P.
PR	28-JUL-1999;	99US-0146222P.	PR	13-NOV-1997;	97US-0065311P.
PR	17-AUG-1999;	99US-0149396P.	PR	25-NOV-1997;	97US-0066770P.
PR	15-SEP-1999;	99WO-US021090.	PR	25-FEB-1998;	98US-0075945P.
PR	15-SEP-1999;	99WO-US021547.	PR	20-MAR-1998;	98US-0078910P.
PR	08-OCT-1999;	99US-0158663P.	PR	28-APR-1998;	98US-0083322P.
PR	30-NOV-1999;	99WO-US028313.	PR	07-MAY-1998;	98US-0084600P.
PR	01-DEC-1999;	99WO-US028301.	PR	28-MAY-1998;	98US-0087106P.
PR	01-DEC-1999;	99WO-US028634.	PR	02-JUN-1998;	98US-0087607P.
PR	16-DEC-1999;	99WO-US030095.	PR	02-JUN-1998;	98US-0087509P.
PR	20-DEC-1999;	99WO-US030911.	PR	02-JUN-1998;	98US-0087759P.
PR	05-JAN-2000;	2000WO-US000219.	PR	03-JUN-1998;	98US-0088271P.
PR	11-FEB-2000;	2000WO-US000376.	PR	04-JUN-1998;	98US-0088021P.
PR	18-FEB-2000;	2000WO-US003565.	PR	04-JUN-1998;	98US-0088025P.
PR	22-FEB-2000;	2000WO-US004341.	PR	04-JUN-1998;	98US-0088026P.
PR	24-FEB-2000;	2000WO-US004414.	PR	04-JUN-1998;	98US-0088028P.
PR	24-FEB-2000;	2000WO-US004914.	PR	04-JUN-1998;	98US-0088029P.
PR	02-MAR-2000;	2000WO-US005004.	PR	04-JUN-1998;	98US-0088030P.
PR	10-MAR-2000;	2000WO-US005841.	PR	04-JUN-1998;	98US-0088033P.
PR	15-MAR-2000;	2000WO-US006319.	PR	05-JUN-1998;	98US-0088326P.
PR	20-MAR-2000;	2000WO-US006884.	PR	05-JUN-1998;	98US-0088167P.
PR	30-MAR-2000;	2000WO-US007377.	PR	05-JUN-1998;	98US-0088202P.
PR	15-MAY-2000;	2000WO-US008439.	PR	05-JUN-1998;	98US-0088212P.
PR	17-MAY-2000;	2000WO-US013358.	PR	05-JUN-1998;	98US-0088217P.
PR	22-MAY-2000;	2000WO-US013705.	PR	09-JUN-1998;	98US-0088655P.
PR	30-MAY-2000;	2000WO-US014042.	PR	10-JUN-1998;	98US-0088734P.
PR	02-JUN-2000;	2000WO-US014941.	PR	10-JUN-1998;	98US-0088738P.
PR	23-JUN-2000;	2000US-0213637P.	PR	10-JUN-1998;	98US-0088742P.
PR	28-JUL-2000;	2000WO-US020710.	PR	10-JUN-1998;	98US-0088810P.
PR	11-AUG-2000;	2000WO-US022031.	PR	10-JUN-1998;	98US-0088824P.
PR	23-AUG-2000;	2000WO-US023522.	PR	10-JUN-1998;	98US-0088826P.
Query Match 100.0%; Score 502; DB 7; Length 104;					98US-0088858P.
Best Local Similarity 100.0%; Pred. No. 1.4e-48;					98US-0088861P.
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					98US-0088876P.
Qy	1	MKLAALGLCVALS	CSAAAFVLSAKPVAQPVAAL	ESAAEAGAGTIANPLGTLNPLKLL	60
Db	1	MKLAALGLCVALS	CSAAAFVLSAKPVAQPVAAL	ESAAEAGAGTIANPLGTLNPLKLL	60
Qy	61	LSSLGIPVNHIEGSKV	AEVGAVKALKALGALTVEG	104	
Db	61	LSSLGIPVNHIEGSKV	AEVGAVKALKALGALTVEG	104	
RESULT 33					
ABO22602					
ID	ABO22602 standard; protein; 104 AA.				
XX	AC ABO22602;				
XX	DT 04-SEP-2003 (first entry)				
XX	DE Human secreted/transmembrane protein PRO1245.				
XX	KW Human; PRO; secreted protein; transmembrane protein; antidiabetic;				
KW	cytostatic; antirheumatic; antiarthritic; antiulcer; neuroprotective;				
KW	antiinflammatory; antibacterial; immunosuppressive; gene therapy;				
KW	diabetes; cancer; rheumatoid arthritis; ulcers;				
KW	amyotrophic lateral sclerosis; inflammatory condition; septic shock.				
OS	Homo sapiens.				
XX	US2003017982-A1.				
XX	PD 23-JAN-2003.				
XX	PF 16-NOV-2001; 2001US-00990441.				
XX	PR 16-JUN-1997; 97US-0049787P.				
PR	17-OCT-1997; 97US-0062250P.				



97USU-0065311P  
97USU-00666770P  
97USU-0075945P  
97USU-0078910P  
97USU-0083322P  
97USU-0084600P  
97USU-0087106P  
97USU-0087607P  
97USU-0087659P  
97USU-0087755P  
97USU-0088021P  
97USU-0088025P  
97USU-0088026P  
97USU-0088028P  
97USU-0088029P  
97USU-0088030P  
97USU-0088033P  
97USU-0088326P  
97USU-0088167P  
97USU-0088202P  
97USU-0088212P  
97USU-0088217P  
97USU-0088655P  
97USU-0088734P  
97USU-0088738P  
97USU-0088743P  
97USU-0088810P  
97USU-0088824P  
97USU-0088826P  
97USU-0088852P  
97USU-0088861P  
97USU-0088876P  
97USU-0089105P  
97USU-0089440P  
97USU-00895112P  
97USU-0089514P  
97USU-0089532P  
97USU-0089533P  
97USU-0089538P  
97USU-0089598P  
97USU-0089947P  
97USU-0089948P  
97USU-0089560P  
97USU-0089653P  
97USU-0089654P  
97USU-0089671P  
97USU-0089908P  
97USU-0089947P  
97USU-0089947P  
97USU-0090444P  
97USU-0090445P  
97USU-0090452P  
97USU-0090453P  
97USU-0090535P  
97USU-0090540P  
97USU-0090542P  
97USU-0090431P  
97USU-0090435P  
97USU-0090444P  
97USU-0090444P  
97USU-0090678P  
97USU-0090678P  
97USU-0090678P  
97USU-0090694P  
97USU-0090694P  
97USU-0090852P  
97USU-0090862P  
97USU-0090862P  
97USU-0090863P  
97USU-0091360P

PR	15-SEP-1999;	99WO-US021547.	PR	24-NOV-2003	(first entry)
PR	08-OCT-1999;	99US-0158663P.	PR	20-NOV-2003	Human secreted/transmembrane protein PRO1245.
PR	30-NOV-1999;	99WO-US028313.	PR	PRO; secreted protein; transmembrane protein;	
PR	01-DEC-1999;	99WO-US028301.	PR	hypertrophy of neonatal heart; angiogenesis;	
PR	01-DEC-1999;	99WO-US028634.	PR	vascular endothelial growth factor; VEGF-stimulated proliferation;	
PR	16-DEC-1999;	99WO-US030095.	PR	endothelial cell; T-lymphocyte proliferation; retinal neuron;	
PR	20-DEC-1999;	99WO-US030911.	PR	c-fos induction; adipocyte cell; chondrocyte differentiation;	
PR	05-JAN-2000;	2000WO-US000219.	PR	pancreatic beta-cell precursor differentiation; gene therapy; tumour;	
PR	06-JAN-2000;	2000WO-US000376.	PR	cancer; human; colon cancer; lung cancer; breast cancer;	
PR	11-FEB-2000;	2000WO-US003565.	PR	rod photoreceptor cell.	
PR	18-FEB-2000;	2000WO-US004341.	OS	Homo sapiens.	
PR	22-FEB-2000;	2000WO-US004414.	XX	US2003059782-A1.	
PR	24-FEB-2000;	2000WO-US004914.	XX	27-MAR-2003.	
PR	24-FEB-2000;	2000WO-US005004.	XX	15-NOV-2001; 2001US-00997628.	
PR	02-MAR-2000;	2000WO-US005841.	XX	16-JUN-1997; 97US-0049787P.	
PR	10-MAR-2000;	2000WO-US006319.	PR	17-OCT-1997; 97US-0062250P.	
PR	15-MAR-2000;	2000WO-US006884.	PR	05-NOV-1997; 97WO-US020069.	
PR	20-MAR-2000;	2000WO-US007377.	PR	12-NOV-1997; 97US-0065186P.	
PR	30-MAR-2000;	2000WO-US008439.	PR	13-NOV-1997; 97US-0065311P.	
PR	15-MAY-2000;	2000WO-US013358.			
PR	17-MAY-2000;	2000WO-US013705.			
PR	22-MAY-2000;	2000WO-US014042.			
PR	30-MAY-2000;	2000WO-US014941.			
PR	02-JUN-2000;	2000WO-US015264.			
PR	23-JUN-2000;	2000US-0213637P.			
PR	28-JUL-2000;	2000WO-US020710.			
PR	11-AUG-2000;	2000WO-US022031.			
Query Match 100.0%; Score 502; DB 7; Length 104;					
Best Local Similarity 100.0%; Pred. No. 1.4e-48; Mismatches 0; Indels 0; Gaps 0;					
Matches 104; Conservative 0; Indels 0; Gaps 0;					
QY	1	MKLAALLGCVALSCSSAAAFVLSAKPVAOPVAALESAAEAGAGTIANPLGTINPLKLL 60			
DB	1	MKLAALLGCVALSCSSAAAFVLSAKPVAOPVAALESAAEAGAGTIANPLGTINPLKLL 60			
QY	61	LSSLGIPVNHLESGSKCVAEILGPQAVGAVKALKALLGALTVPFG 104			
DB	61	LSSLGIPVNHLESGSKCVAEILGPQAVGAVKALKALLGALTVPFG 104			
RESULT 35					
ADA39390					
ID	ADA39390 standard; protein; 104 AA.				
XX	ADA39390;				
AC					
XX	20-NOV-2003 (first entry)				
DT	Human secreted/transmembrane protein PRO1245.				
XX	PRO; secreted protein; transmembrane protein;				
XX	hypertrophy of neonatal heart; angiogenesis;				
KW	vascular endothelial growth factor; VEGF-stimulated proliferation;				
KW	endothelial cell; T-lymphocyte proliferation; retinal neuron;				
KW	c-fos induction; adipocyte cell; chondrocyte differentiation;				
KW	pancreatic beta-cell precursor differentiation; gene therapy; tumour;				
KW	cancer; human; colon cancer; lung cancer; breast cancer;				
KW	rod photoreceptor cell.				
XX					
OS	Homo sapiens.				
XX					
XX	US2003059782-A1.				
XX					
XX	27-MAR-2003.				
XX					
XX	15-NOV-2001; 2001US-00997628.				
XX					
PR	16-JUN-1997; 97US-0049787P.				
PR	17-OCT-1997; 97US-0062250P.				
PR	05-NOV-1997; 97WO-US020069.				
PR	12-NOV-1997; 97US-0065186P.				
PR	13-NOV-1997; 97US-0065311P.				



[illegible]



PR 09-JUL-1998; 98US-0092182P.  
PR 10-JUL-1998; 98US-0092472P.  
PR 20-JUL-1998; 98US-0093333P.  
PR 30-JUL-1998; 98US-0094651P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 04-AUG-1998; 98US-0095283P.  
PR 04-AUG-1998; 98US-0095301P.  
PR 04-AUG-1998; 98US-0095302P.  
PR 04-AUG-1998; 98US-0095318P.  
PR 04-AUG-1998; 98US-0095321P.  
PR 04-AUG-1998; 98US-0095325P.  
PR 04-AUG-1998; 98US-0095915P.  
PR 10-AUG-1998; 98US-0095923P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 11-AUG-1998; 98US-0096143P.  
PR 11-AUG-1998; 98US-0096146P.  
PR 12-AUG-1998; 98US-0096329P.  
PR 12-AUG-1998; 98US-0096329P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096768P.  
PR 17-AUG-1998; 98US-0096773P.  
PR 17-AUG-1998; 98US-0096791P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096894P.  
PR 17-AUG-1998; 98US-0096895P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096894P.  
PR 18-AUG-1998; 98US-0096950P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0096960P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 19-AUG-1998; 98US-0097141P.  
PR 20-AUG-1998; 98US-0097218P.  
PR 24-AUG-1998; 98US-0097661P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0097978P.  
PR 26-AUG-1998; 98US-0097979P.  
PR 26-AUG-1998; 98US-0097986P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 31-AUG-1998; 98US-0098525P.  
PR 16-SEP-1998; 98US-0100634P.  
PR 16-SEP-1998; 98US-0100634P.  
PR 17-SEP-1998; 98US-0100858P.  
PR 17-SEP-1998; 98US-0100858P.  
PR 07-OCT-1998; 98US-0101943P.  
PR 01-DEC-1998; 98US-0113236P.  
PR 22-DEC-1998; 98US-0113236P.  
PR 05-JAN-1999; 98US-0113236P.  
PR 08-MAR-1999; 98US-0123957P.  
PR 12-MAR-1999; 98US-0123957P.  
PR 02-JUN-1999; 98US-0141225P.  
PR 23-JUN-1999; 98US-0141037P.  
PR 07-JUL-1999; 98US-0143048P.  
PR 20-JUL-1999; 98US-0144758P.  
PR 26-JUL-1999; 98US-0145698P.  
PR 28-JUL-1999; 98US-0146222P.  
PR 17-AUG-1999; 98US-0149396P.  
PR 15-SEP-1999; 98US-0201090P.  
PR 15-SEP-1999; 98US-0201547P.  
PR 08-OCT-1999; 98US-0158663P.  
PR 30-NOV-1999; 98US-0202831P.  
PR 01-DEC-1999; 98US-0202831P.  
PR 01-DEC-1999; 98US-0202831P.  
PR 16-DEC-1999; 98US-0202831P.  
PR 20-DEC-1999; 98US-0202831P.  
PR 05-JAN-2000; 98US-0202831P.  
PR 06-JAN-2000; 98US-0202831P.  
PR 11-FEB-2000; 98US-0202831P.

PR 18-FEB-2000; 2000WO-US004341.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 15-MAY-2000; 2000WO-US013358.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 23-JUN-2000; 2000US-0213637P.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
  
Query Match 100.0%; Score 502; DB 7; Length 104;  
Best Local Similarity 100.0%; Pred. No. 1.4e-48;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MKLAALGLCVALLSCSSAAAFVLSAKPVAQPVAALESAAAGAGTIANPLGTLPKLL 60  
Db 1 MKLAALGLCVALLSCSSAAAFVLSAKPVAQPVAALESAAAGAGTIANPLGTLPKLL 60  
  
QY 61 LSSGIPVNHLEIGSQKVAELGPQAVGAVKALKALIGALTVEG 104  
Db 61 LSSGIPVNHLEIGSQKVAELGPQAVGAVKALKALIGALTVEG 104  
  
RESULT 38  
ADC52522  
ID ADC52522 standard; protein; 104 AA.  
XX ADC52522;  
AC ADC52522;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Human PRO polypeptide #127.  
XX  
KW Human; PRO; pancreatic beta-cell precursor cell; pancreatic beta-cell;  
KW insulin deficiency; diabetes mellitus; haemoglobin-associated disorder;  
KW thalassemia; endothelial cell growth; cancer; cystic renal dysplasia;  
KW polycystic kidney disease; renal tumour; antidiabetic; antianemic;  
KW cytostatic; cardiac; vulnary; antinflammatory; anorectic.  
XX  
OS Homo sapiens.  
XX  
US2003045463-A1.  
XX  
PD 06-MAR-2003.  
XX  
PF 16-NOV-2001; 2001US-00990437.  
XX  
PR 16-JUN-1997; 97US-0049787P.  
PR 17-OCT-1997; 97US-0062250P.  
PR 03-NOV-1997; 97US-0062008P.  
PR 12-NOV-1997; 97US-0065186P.  
PR 13-NOV-1997; 97US-0065311P.  
PR 24-NOV-1997; 97US-0066770P.  
PR 25-FEB-1998; 98US-0075945P.  
PR 20-MAR-1998; 98US-0078910P.  
PR 28-APR-1998; 98US-0083322P.  
PR 07-MAY-1998; 98US-0087106P.  
PR 28-MAY-1998; 98US-0087106P.  
PR 02-JUN-1998; 98US-0087607P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 03-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088021P.  
PR 04-JUN-1998; 98US-0088025P.





PR	04-JUN-1998	98US-008026P
PR	04-JUN-1998	98US-008028P
PR	04-JUN-1998	98US-008029P
PR	04-JUN-1998	98US-008030P
PR	04-JUN-1998	98US-008033P
PR	04-JUN-1998	98US-008032P
PR	05-JUN-1998	98US-0080167P
PR	05-JUN-1998	98US-008020P
PR	05-JUN-1998	98US-008021P
PR	05-JUN-1998	98US-0080217P
PR	05-JUN-1998	98US-008065P
PR	09-JUN-1998	98US-0080734P
PR	10-JUN-1998	98US-0080738P
PR	10-JUN-1998	98US-0080742P
PR	10-JUN-1998	98US-0080840P
PR	10-JUN-1998	98US-0080824P
PR	10-JUN-1998	98US-0080826P
PR	11-JUN-1998	98US-00808958P
PR	11-JUN-1998	98US-00808961P
PR	11-JUN-1998	98US-0080876P
PR	12-JUN-1998	98US-0080105P
PR	16-JUN-1998	98US-0080944P
PR	16-JUN-1998	98US-00809512P
PR	16-JUN-1998	98US-00809532P
PR	17-JUN-1998	98US-00809534P
PR	17-JUN-1998	98US-00809538P
PR	17-JUN-1998	98US-00809598P
PR	17-JUN-1998	98US-00809599P
PR	17-JUN-1998	98US-00809600P
PR	17-JUN-1998	98US-00809653P
PR	18-JUN-1998	98US-00809801P
PR	18-JUN-1998	98US-0080907P
PR	18-JUN-1998	98US-0080908P
PR	19-JUN-1998	98US-0080947P
PR	19-JUN-1998	98US-0080948P
PR	19-JUN-1998	98US-0080952P
PR	22-JUN-1998	98US-0090246P
PR	22-JUN-1998	98US-0090252P
PR	22-JUN-1998	98US-0090254P
PR	23-JUN-1998	98US-0090349P
PR	23-JUN-1998	98US-0090335P
PR	24-JUN-1998	98US-0090429P
PR	24-JUN-1998	98US-0090411P
PR	24-JUN-1998	98US-0080435P
PR	24-JUN-1998	98US-0090444P
PR	24-JUN-1998	98US-0090445P
PR	24-JUN-1998	98US-0090472P
PR	25-JUN-1998	98US-0090533P
PR	25-JUN-1998	98US-0090540P
PR	25-JUN-1998	98US-0090542P
PR	25-JUN-1998	98US-0090557P
PR	25-JUN-1998	98US-0090567P
PR	25-JUN-1998	98US-0090678P
PR	25-JUN-1998	98US-0090690P
PR	25-JUN-1998	98US-0090694P
PR	25-JUN-1998	98US-0091519P
PR	02-JUL-1998	98US-0091626P
PR	02-JUL-1998	98US-0091633P
PR	02-JUL-1998	98US-0091628P
PR	02-JUL-1998	98US-0091646P
PR	02-JUL-1998	98US-0091673P
PR	07-JUL-1998	98US-0091978P
PR	07-JUL-1998	98US-0091982P
PR	09-JUL-1998	98US-0092182P
PR	10-JUL-1998	98US-0092472P
PR	20-JUL-1998	98US-0093319P

```
PR 30-JUL-1998; 98US-0094651P.
PR 04-AUG-1998; 98US-0095282P.
PR 04-AUG-1998; 98US-0095285P.
PR 04-AUG-1998; 98US-0095301P.
PR 04-AUG-1998; 98US-0095302P.
PR 04-AUG-1998; 98US-0095318P.
PR 04-AUG-1998; 98US-0095321P.
PR 04-AUG-1998; 98US-0095325P.
PR 10-AUG-1998; 98US-0095916P.
PR 10-AUG-1998; 98US-0095929P.
PR 10-AUG-1998; 98US-0096122P.
PR 11-AUG-1998; 98US-0096143P.
PR 11-AUG-1998; 98US-0096146P.
PR 12-AUG-1998; 98US-0096329P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096768P.
PR 17-AUG-1998; 98US-0096773P.
PR 17-AUG-1998; 98US-0096791P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096894P.
PR 17-AUG-1998; 98US-0096895P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096849P.
PR 18-AUG-1998; 98US-0096850P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0096960P.
PR 18-AUG-1998; 98US-0097022P.
PR 19-AUG-1998; 98US-0097141P.
PR 20-AUG-1998; 98US-0097218P.
PR 20-AUG-1998; 98US-0097661P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0097978P.
PR 26-AUG-1998; 98US-0097979P.
PR 26-AUG-1998; 98US-0097986P.
PR 26-AUG-1998; 98US-0098014P.
PR 31-AUG-1998; 98US-0098525P.
PR 16-SEP-1998; 98US-0100634P.
PR 17-SEP-1998; 98US-0100858P.
PR 17-SEP-1998; 98US-0101943P.
PR 07-OCT-1998; 98WO-US021141.
PR 01-DEC-1998; 98WO-US025108.
PR 22-DEC-1998; 98US-0113296P.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US0005028.
PR 12-MAR-1999; 99US-01233957P.
PR 02-JUN-1999; 99WO-US012252.
PR 23-JUN-1999; 99US-0141037P.
PR 07-JUL-1999; 99US-0143048P.
PR 20-JUL-1999; 99US-0144758P.
PR 26-JUL-1999; 99US-0145698P.
PR 28-JUL-1999; 99US-0146222P.
PR 17-AUG-1999; 99US-0149396P.
PR 15-SEP-1999; 99WO-US021090.
PR 15-SEP-1999; 99WO-US021547.
PR 08-OCT-1999; 99US-0158863P.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028634.
PR 16-DEC-1999; 99WO-US030911.
PR 20-DEC-1999; 99WO-US030911.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US000365.
PR 18-FEB-2000; 2000WO-US0004341.
PR 22-FEB-2000; 2000WO-US0004414.
PR 24-FEB-2000; 2000WO-US0004914.

PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 15-MAR-2000; 2000WO-US006884.
PR 20-MAR-2000; 2000WO-US007377.
PR 30-MAR-2000; 2000WO-US008439.
PR 15-MAY-2000; 2000WO-US013358.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 23-JUN-2000; 2000US-0213637P.

Query Match 100.0%; Score 502; DB 7; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.4e-48;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKLAALLGLCVALLSCSSAAFLVGSAPVQAQVAALESAAEAGAGTLANPLGTINPLKLL 60
Db 1 MKLAALLGLCVALLSCSSAAFLVGSAPVQAQVAALESAAEAGAGTLANPLGTINPLKLL 60
Qy 61 LSSLGIPVNHLEIGSQKVAELGPOAVGAVKALKALLGALTVEG 104
Db 61 LSSLGIPVNHLEIGSQKVAELGPOAVGAVKALKALLGALTVEG 104

RESULT 40
ADC56541
ID ADC56541 standard; protein; 104 AA.
XX AC ADC56541;
XX DT 18-DEC-2003 (first entry)
XX DE Human PRO polypeptide #127.
XX KW Human; PRO; pancreatic beta-cell precursor cell; pancreatic beta-cell;
XX KW insulin deficiency; diabetes mellitus; haemoglobin-associated disorder;
XX KW thalassaemia; endothelial cell growth; cancer; cystic renal dysplasia;
XX KW polycystic kidney disease; renal tumour; antidiabetic; antianaemic;
XX KW cytostatic; cardiant; vulnery; antinflammatory; anorectic.
XX OS Homo sapiens.
XX US2003064375-A1.
XX PD 03-APR-2003.
XX PF 15-NOV-2001; 2001US-00997857.
XX 16-JUN-1997; 97US-0049787P.
XX 17-OCT-1997; 97US-0062250P.
XX 05-NOV-1997; 97WO-US020069.
XX 12-NOV-1997; 97US-0065186P.
XX 13-NOV-1997; 97US-0065311P.
XX 24-NOV-1997; 97US-0066770P.
XX 25-FEB-1998; 98US-0075945P.
XX 20-MAR-1998; 98US-0078910P.
XX 28-APR-1998; 98US-0083322P.
XX 07-MAY-1998; 98US-0084600P.
XX 28-MAY-1998; 98US-0087106P.
XX 02-JUN-1998; 98US-0087607P.
XX 02-JUN-1998; 98US-0087609P.
XX 02-JUN-1998; 98US-0087759P.
XX 03-JUN-1998; 98US-0087827P.
XX 04-JUN-1998; 98US-0088021P.
XX 04-JUN-1998; 98US-0088025P.
XX 04-JUN-1998; 98US-0088026P.
XX 04-JUN-1998; 98US-0088028P.
XX 04-JUN-1998; 98US-0088029P.
XX 04-JUN-1998; 98US-0088030P.
XX 04-JUN-1998; 98US-0088033P.
XX 04-JUN-1998; 98US-0088326P.
```

PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088734P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088742P.  
PR 10-JUN-1998; 98US-0088810P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088846P.  
PR 11-JUN-1998; 98US-0088856P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 12-JUN-1998; 98US-0089440P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 16-JUN-1998; 98US-0089532P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089599P.  
PR 17-JUN-1998; 98US-0089600P.  
PR 17-JUN-1998; 98US-0089651P.  
PR 18-JUN-1998; 98US-0089801P.  
PR 18-JUN-1998; 98US-0089907P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089947P.  
PR 19-JUN-1998; 98US-0089948P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 23-JUN-1998; 98US-0090349P.  
PR 23-JUN-1998; 98US-0090355P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090431P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090445P.  
PR 24-JUN-1998; 98US-0090472P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 24-JUN-1998; 98US-0090542P.  
PR 24-JUN-1998; 98US-0090557P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 01-JUL-1998; 98US-0091360P.  
PR 01-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091519P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091633P.  
PR 02-JUL-1998; 98US-0091646P.  
PR 02-JUL-1998; 98US-0091673P.  
PR 02-JUL-1998; 98US-0091978P.  
PR 07-JUL-1998; 98US-0091982P.  
PR 07-JUL-1998; 98US-0092182P.  
PR 10-JUL-1998; 98US-0092472P.  
PR 10-JUL-1998; 98US-0093339P.  
PR 30-JUL-1998; 98US-0094511P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 04-AUG-1998; 98US-0095285P.  
PR 04-AUG-1998; 98US-0095301P.  
PR 04-AUG-1998; 98US-0095302P.  
PR 04-AUG-1998; 98US-0095318P.

PR 04-AUG-1998; 98US-0095321P.  
PR 04-AUG-1998; 98US-0095325P.  
PR 10-AUG-1998; 98US-0095916P.  
PR 10-AUG-1998; 98US-0095929P.  
PR 10-AUG-1998; 98US-0096012P.  
PR 11-AUG-1998; 98US-0096143P.  
PR 11-AUG-1998; 98US-0096146P.  
PR 12-AUG-1998; 98US-0096329P.  
PR 12-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096768P.  
PR 17-AUG-1998; 98US-0096773P.  
PR 17-AUG-1998; 98US-0096791P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096894P.  
PR 17-AUG-1998; 98US-0096895P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096950P.  
PR 18-AUG-1998; 98US-0096959P.  
PR 18-AUG-1998; 98US-0096960P.  
PR 18-AUG-1998; 98US-0097022P.  
PR 18-AUG-1998; 98US-0097141P.  
PR 20-AUG-1998; 98US-0097218P.  
PR 24-AUG-1998; 98US-0097661P.  
PR 26-AUG-1998; 98US-0097952P.  
PR 26-AUG-1998; 98US-0097954P.  
PR 26-AUG-1998; 98US-0097955P.  
PR 26-AUG-1998; 98US-0097971P.  
PR 26-AUG-1998; 98US-0097974P.  
PR 26-AUG-1998; 98US-0097978P.  
PR 26-AUG-1998; 98US-0097979P.  
PR 26-AUG-1998; 98US-0097986P.  
PR 26-AUG-1998; 98US-0098014P.  
PR 31-AUG-1998; 98US-0098525P.  
PR 16-SEP-1998; 98US-0100634P.  
PR 16-SEP-1998; 98US-0100858P.  
PR 17-SEP-1998; 98US-0100858P.  
PR 17-SEP-1998; 98US-0100858P.  
PR 07-OCT-1998; 98US-0101141P.  
PR 01-DEC-1998; 98US-01025108.  
PR 22-DEC-1998; 98US-0113296P.  
PR 05-JAN-1999; 99US-0143396P.  
PR 08-MAR-1999; 99US-0123957P.  
PR 12-MAR-1999; 99US-0123957P.  
PR 02-JUN-1999; 99US-0141037P.  
PR 23-JUN-1999; 99US-0141037P.  
PR 07-JUL-1999; 99US-0143048P.  
PR 20-JUL-1999; 99US-0143048P.  
PR 26-JUL-1999; 99US-0145686P.  
PR 28-JUL-1999; 99US-0146222P.  
PR 17-AUG-1999; 99US-0149396P.  
PR 15-SEP-1999; 99US-0149396P.  
PR 15-SEP-1999; 99US-0149396P.  
PR 08-OCT-1999; 99US-0158663P.  
PR 30-NOV-1999; 99US-0158663P.  
PR 01-DEC-1999; 99US-0158663P.  
PR 01-DEC-1999; 99US-0158663P.  
PR 16-DEC-1999; 99US-0158663P.  
PR 20-DEC-1999; 99US-0158663P.  
PR 05-JAN-2000; 2000US-000219.  
PR 06-JAN-2000; 2000US-000376.  
PR 11-FEB-2000; 2000US-000376.  
PR 18-FEB-2000; 2000US-000434.  
PR 22-FEB-2000; 2000US-000434.  
PR 24-FEB-2000; 2000US-0004914.  
PR 24-FEB-2000; 2000US-0005004.  
PR 02-MAR-2000; 2000US-0005841.  
PR 10-MAR-2000; 2000US-0006319.  
PR 15-MAR-2000; 2000US-0006884.  
PR 20-MAR-2000; 2000US-0007377.  
PR 30-MAR-2000; 2000US-0008439.

```
PR 15-MAY-2000; 2000WO-US013358.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 23-JUN-2000; 2000US-0213637P.
PR 28-JUL-2000; 2000WO-US020710.
PR 11-AUG-2000; 2000WO-US022031.
PR 23-AUG-2000; 2000WO-US023522.

Query Match      100.0%; Score 502; DB 7; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.4e-48;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKLAALLGLCVLSCSSAAAFLVGSAPVQPVAALESAAEAGAGTIANPLGTLNPLKLL 60
Db 1 MKLAALLGLCVLSCSSAAAFLVGSAPVQPVAALESAAEAGAGTIANPLGTLNPLKLL 60

Qy 61 LSSLGIPVNHLEGSQKVAELGPQAVGAVKALKALLGALTVEG 104
Db 61 LSSLGIPVNHLEGSQKVAELGPQAVGAVKALKALLGALTVEG 104
```

Search completed: April 5, 2004, 14:34:07  
Job time : 58 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 5, 2004, 14:11:13 ; Search time 17 seconds  
(without alignments)

318,547 Million cell updates/sec

Title: US-09-997-428-408

Perfect score: 502

Sequence: 1 MKLAALLGLCVSLSSGAA.....QAVGAVKALKALLGALTTFVG 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	497	99.0	104	UGR2_HUMAN	Q96qr1 homo sapien
2	250	49.8	104	UGR2_MOUSE	Q920d7 mus musculus
3	160	31.9	93	UGR1_HUMAN	Q96p11 homo sapien
4	135	26.9	139	UGR1_MOUSE	Q920h1 mus musculus
5	83	16.5	1327	TNK1_HUMAN	O95271 homo sapien
6	77	15.3	335	TRD2_STRCO	Q924w9 streptomyce
7	76	15.1	732	COPA_HELPF	O32619 helicobacte
8	75	14.9	362	MDC1_HUMAN	Q9h1k6 homo sapien
9	75	14.9	362	MDC1_MOUSE	Q9ere8 mus musculus
10	74	14.7	779	SNL1_MOUSE	Q60670 mus musculus
11	73.5	14.6	626	DXS_WIGBR	Qad357 wiggleswort
12	73.5	14.6	776	SNL1_RAT	Q9riu5 rattus norv
13	73.5	14.6	882	SYA_THETH	P74941 thermus th
14	72	14.3	397	SUCC_AGRYS	Q8uc60 agrobacteri
15	71	14.1	428	FXB2_MOUSE	Q64733 mus musculus
16	71	14.1	480	DNAA_RHIME	P35890 rhizobium m
17	69	13.7	124	RL7_RALSO	Q9xuz7 ralstonia s
18	68.5	13.6	377	PROB_RIFLO	Q9g4t9 bifidobacte
19	68	13.5	234	FEPD_SCOLI	P31876 escherichia
20	68	13.5	569	SIUF_MOUSE	Q920g3 mus musculus
21	67.5	13.4	2035	Y213_HUMAN	Q92508 homo sapien
22	67	13.3	351	GCP_MYCLE	P37969 mycobacteri
23	67	13.3	464	FUNC_PSESM	Q885v0 pseudomonas
24	66.5	13.2	341	TAZR_RAT	P34978 rattus norv
25	66.5	13.2	465	FXD1_HUMAN	Q16676 homo sapien
26	66	13.1	118	NLID_BRAOL	Q43304 brassica ol
27	66	13.1	121	MP70_WYCKA	Q49614 mycobacteri
28	66	13.1	239	CLRC_IDEDE	P60000 ideonella d
29	66	13.1	527	FTB_MOUSE	P17225 mus musculus
30	66	13.1	540	LEUI_PROMM	Q7tut5 prochloroco
31	66	13.1	585	NUP1_RAT	P70581 rattus norv
32	65.5	13.0	257	EUTC_RHOER	Q59782 rhodococcu
33	65.5	13.0	336	COBT_RHILO	Q9Bkn9 rhizobium l

RESULT 1

#### ALIGNMENTS

34	65.5	13.0	491	1	Y084_MYCTU
35	65	12.9	344	1	TRD2_RALSO
36	65	12.9	399	1	YM96_SYNEL
37	65	12.9	462	1	A2AC_HUMAN
38	65	12.9	693	1	RECG_ECOLI
39	65	12.9	4377	1	ANK3_HUMAN
40	64.5	12.8	184	1	CYCM_BRAJA
41	64.5	12.8	320	1	Y678_METJA
42	64.5	12.8	750	1	ELS_CHICK
43	64.5	12.8	905	1	NUOG_PSEAE
44	64.5	12.8	1023	1	RT11_ACIPL
45	64.5	12.8	1023	1	RT12_ACIPL
46	64.5	12.8	1159	1	RPOC_PORCN
47	64.5	12.8	1608	1	HLYA_SERMA
48	64.5	12.8	2541	1	TIN1_HUMAN
49	64	12.7	246	1	HIS4_METWA
50	64	12.7	327	1	THI4_ASPOR
51	64	12.7	426	1	Y200_ANASP
52	64	12.7	487	1	DNAA_AGR75
53	64	12.7	655	1	YAGF_ECOLI
54	64	12.7	693	1	RECG_ECO57
55	63.5	12.6	82	1	C551_PSEME
56	63.5	12.6	130	1	F14A_HUMAN
57	63.5	12.6	130	1	RL7_MYCLE
58	63.5	12.6	230	1	PLGH_XANCP
59	63.5	12.6	345	1	TRPD_AERPE
60	63.5	12.6	394	1	YGAY_ECO57
61	63.5	12.6	394	1	YGAY_ECOLI
62	63.5	12.6	397	1	Y435_PROMA
63	63.5	12.6	407	1	AROC_MYCLE
64	63.5	12.6	432	1	YF10_MYCTU
65	63.5	12.6	541	1	Y4NM_RHISN
66	63.5	12.6	548	1	MEBA_PSEFL
67	63.5	12.6	1048	1	P100_HCWVA
68	63	12.5	434	1	YU18_MYCTU
69	63	12.5	435	1	YU21_MYCTU
70	63	12.5	517	1	6PGD_CANAL
71	63	12.5	932	1	PMS1_HUMAN
72	63	12.5	1062	1	NAL2_HUMAN
73	62.5	12.5	267	1	THIG_DEIRA
74	62.5	12.5	361	1	COBT_MYCTU
75	62.5	12.5	511	1	NADB_CAUCR
76	62.5	12.5	560	1	EFS_MOUSE
77	62.5	12.5	760	1	CO2_MOUSE
78	62	12.4	128	1	RL7_CORGL
79	62	12.4	199	1	IL11_RAT
80	62	12.4	242	1	HIS4_HALN1
81	62	12.4	257	1	MOAE_KLEAE
82	62	12.4	299	1	G3P_DICDI
83	62	12.4	382	1	YN28_MYCTU
84	62	12.4	388	1	CHSD_PHANI
85	62	12.4	407	1	PCAB_PSEFU
86	62	12.4	446	1	MURF_STRTO
87	62	12.4	620	1	SNP_CQTUA
88	62	12.4	660	1	REP2_HUMAN
89	62	12.4	668	1	TKT_XANFL
90	62	12.4	775	1	YTX1_XENLA
91	62	12.4	829	1	E74A_DROME
92	62	12.4	883	1	E74B_DROME
93	62	12.4	954	1	GCSP_AGR75
94	61.5	12.3	124	1	RL7_BRUME
95	61.5	12.3	217	1	RNH2_AGR75
96	61.5	12.3	306	1	DDL_HAEIN
97	61.5	12.3	318	1	KIPF_XANCP
98	61.5	12.3	344	1	NAGK_HUMAN
99	61.5	12.3	355	1	TRPD_AZOBR
100	61.5	12.3	367	1	NK61_HUMAN

UGR2\_HUMAN  
ID UGR2\_HUMAN STANDARD; PRT; 104 AA.  
AC Q96QL1; Q96PL0;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Uteroglobin-related protein 2 precursor (Cytokine HIN-1) (High in  
DE normal-1) (Secretogloblin family 3A member 1).  
GN SCGB3A1 OR UGRP2 OR HIN1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21396515; PubMed=11481438;  
RA Krop I.E., Sgroi D., Porter D.A., Lunetta K.L., LeVangie R., Seth P.,  
RA Kaelin C.M., Rhei E., Bosenberg M., Schnitt S., Marks J.R., Pagon Z.,  
RA Belina D., Razumovic J., Polyak K.;  
RT "HIN-1, a putative cytokine highly expressed in normal but not  
RT cancerous mammary epithelial cells.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9796-9801 (2001). 08/14/01  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21539178; PubMed=11682631;  
RA Niimi T., Keck-Waggoner C.L., Popescu N.C., Zhou Y., Levitt R.C.,  
RA Kimura S.;  
RT "UGRP1, a uteroglobin/clara cell secretory protein-related protein, is  
RT a novel lung-enriched downstream target gene for the T/EBP/NKX2.1  
RT homeodomain transcription factor.";  
RL Mol. Endocrinol. 15:2021-2036 (2001).  
CC -!- FUNCTION: Potential growth inhibitory cytokine.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Highly expressed in breast tissues. Absent in  
CC breast cancer cell lines.  
CC -!- SIMILARITY: Belongs to the uteroglobin family. UGRP subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF040564; AAK82942.1; -;  
CC EMBL; AF313458; AAL26217.1; -;  
CC Genew; HGNC:18384; SCGB3A1.  
CC MIM; 606500; -;  
CC GO; GO:000576; C:extracellular; NAS.  
CC GO; GO:0005125; F:cytokine activity; NAS.  
CC GO; GO:0030308; P:negative regulation of cell growth; NAS.  
CC GO; GO:0042127; P:regulation of cell proliferation; NAS.  
KW Cytokine; Signal.  
FT SIGNAL 1 20 POTENTIAL.  
FT CHAIN 21 104 UTEROGLIBIN-RELATED PROTEIN 2.  
FT CONFLICT 19 19 R -> A (IN REF. 2).  
FT SEQUENCE 104 AA; 10185 MW; 1083873C8FAE8015 CRC64;  
SQ  
Query Match 99.0%; Score 497; DB 1; Length 104;  
Best Local Similarity 99.0%; Pred. No. 7.2e-38;  
Matches 103; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MKLALLGLCVALSCSAAFLVGSAPQVPAALAESAAEAGAGTLANPLGTINPLKLL 60  
DB 1 MKLALLGLCVALSCSAAFLVGSAPQVPAALAESAAEAGAGTLANPLGTINPLKLL 60  
QY 61 LSSLGIPVNHLEIGSKQKVAELGPQAVGAVKALKALIGALTIVFG 104  
DB 61 LSSLGIPVNHLEIGSKQKVAELGPQAVGAVKALKALIGALTIVFG 104  
RESULT 2

UGR2\_MOUSE  
ID UGR2\_MOUSE STANDARD; PRT; 104 AA.  
AC Q920D7;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Uteroglobin-related protein 2 precursor (Cytokine HIN-1) (High in  
DE normal-1) (Secretogloblin family 3A member 1).  
GN SCGB3A1 OR UGRP2 OR HIN1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21396515; PubMed=11481438;  
RA Krop I.E., Sgroi D., Porter D.A., Lunetta K.L., LeVangie R., Seth P.,  
RA Kaelin C.M., Rhei E., Bosenberg M., Schnitt S., Marks J.R., Pagon Z.,  
RA Belina D., Razumovic J., Polyak K.;  
RT "HIN-1, a putative cytokine highly expressed in normal but not  
RT cancerous mammary epithelial cells.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9796-9801 (2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21539178; PubMed=11682631;  
RA Niimi T., Keck-Waggoner C.L., Popescu N.C., Zhou Y., Levitt R.C.,  
RA Kimura S.;  
RT "UGRP1, a uteroglobin/clara cell secretory protein-related protein, is  
RT a novel lung-enriched downstream target gene for the T/EBP/NKX2.1  
RT homeodomain transcription factor.";  
RL Mol. Endocrinol. 15:2021-2036 (2001).  
CC -!- FUNCTION: Potential growth inhibitory cytokine.  
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).  
CC -!- SIMILARITY: Belongs to the uteroglobin family. UGRP subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF313456; AAL26216.1; -;  
CC MGD; MGI:1915912; Scgb3a1.  
CC SIGNAL 1 21 POTENTIAL.  
CC CHAIN 22 104 UTEROGLIBIN-RELATED PROTEIN 2.  
CC SEQUENCE 104 AA; 10591 MW; D62F0E601FB57A6D CRC64;  
SQ  
Query Match 49.8%; Score 250; DB 1; Length 104;  
Best Local Similarity 57.0%; Pred. No. 8.2e-16;  
Matches 61; Conservative 14; Mismatches 26; Indels 6; Gaps 4;  
QY 1 MKLALLGLCVALSCSAAFLVGSAPQVPAALAESAAEAGAGTLANPLGTINPL 57  
DB 1 MKLTTTFLVLCVALLSDSGVAFPMDSLAKPAVEPVAALAPAAEAVAGVPSIPLSHLAIL 60  
QY 58 KLLSLGIPVNHLEIGSKQKVAELGPQAVGAVKALKALIGALTIVFG 104  
DB 61 RFLASNGIPLDLIEGRKRCVTELGPEAVGAV---KSLGLVLTWFG 104  
RESULT 3  
UGR1\_HUMAN  
ID UGR1\_HUMAN STANDARD; PRT; 93 AA.  
AC Q96PL1;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Uteroglobin-related protein 1 precursor (Secretogloblin family 3A  
DE member 2).  
GN SCGB3A2 OR UGRP1.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21539178; PubMed=11682631;  
RA Kimura S.;  
RA Niimi T.; Kekk-Waggoner C.L.; Popescu N.C.; Zhou Y.; Levitt R.C.;  
RT "UGRP1, a uteroglobin/clara cell secretory protein-related protein, is  
a novel lung-enriched downstream target gene for the T/EBP/NKX2.1  
homeodomain transcription factor.";  
RL Mol. Endocrinol. 15:2021-2036(2001).  
[2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;  
RA Klausner R.D.; Collins F.S.; Wagner L.; Shenmen C.M.; Schuler G.D.;  
RA Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;  
RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Heide F.;  
RA Datchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;  
RA Stapleton M.; Soares M.B.; Bonaldo M.P.; Casavant T.L.; Scheetz T.E.;  
RA Brownstein M.J.; Udwin T.B.; Toshiyuki S.; Carninci P.; Prange C.;  
RA Raha S.S.; Loquellano N.A.; Peters G.J.; Abramson R.D.; Mullany S.J.;  
RA Bobak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;  
RA Richards S.; Woxley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.;  
RA Villalon D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;  
RA Fahey J.; Helton E.; Kettman M.; Madan A.; Rodriguez S.; Sanchez A.;  
RA Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;  
RA Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;  
RA Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.M.;  
RA Butlerfield Y.S.N.; Krzywinski M.I.; Skaleka U.; Smailus D.E.;  
RA Schnerch A.; Schein J.E.; Jones S.J.M.; Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC  
CC -!- SUBUNIT: Homodimer.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Highly expressed in lung.  
CC -!- SIMILARITY: Belongs to the uteroglobin family. UGRP subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF313455; AAL26215.1; -;  
CC EMBL; BC024232; AAL24232.1; -;  
CC Genbank; HGNC:18391; SCGB3A2.  
CC MIM; 606531; -;  
KW SIGNAL.  
FT CHAIN 1 21 POTENTIAL.  
FT SIGNAL 22 93 UTEROGLOBIN-RELATED PROTEIN 1.  
FT CHAIN 22 93  
SQ SEQUENCE 93 AA; 10161 MW; FBD4BFAC2BF3718 CRC64;  
Query Match 31.9%; Score 160; DB 1; Length 93;  
Best Local Similarity 43.6%; Pred. No. 8.1e-08;  
Matches 44; Conservative 9; Mismatches 38; Indels 10; Gaps 2;  
QY 1 MKLAALGLCVALS-  
Db 1 MKLVTFILVTISCVSATAFLINKVPLPVDKLAFL-----PLDNLFPMDPLKL 51  
QY 60 LLSLIGIPVNHLEGSQKVAELGPQAVGAVKALKALLGAL 100  
Db 52 LLKLTGISVHEVLGRKCNVELGPEASEAVKKLEALSIL 92

RESULT 4

UGRP1\_MOUSE  
ID UGRL\_MOUSE STANDARD; PRT; 139 AA.  
AC Q920H1; Q920H2; Q920H3;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Uteroglobin-related protein 1 precursor (Secretoglobin family 3A  
member 2).  
DE member 2).  
GN SCGB3A2 OR UGRP1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RN SEQUENCE FROM N.A. (ISOFORMS A; B AND C).  
RP TISSUE=Lung;  
RX MEDLINE=21539178; PubMed=11682631;  
RA Niimi T.; Kekk-Waggoner C.L.; Popescu N.C.; Zhou Y.; Levitt R.C.;  
RA Kimura S.;  
RT "UGRP1, a uteroglobin/clara cell secretory protein-related protein, is  
a novel lung-enriched downstream target gene for the T/EBP/NKX2.1  
homeodomain transcription factor.";  
RL Mol. Endocrinol. 15:2021-2036(2001).  
CC -!- SUBUNIT: Homodimer.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Name=C;  
CC IsoId=Q920H1-1; Sequence=Displayed;  
CC Name=A;  
CC IsoId=Q920H1-2; Sequence=VSP\_006727, VSP\_006728;  
CC Name=B;  
CC IsoId=Q920H1-3; Sequence=VSP\_006726;  
CC -!- TISSUE SPECIFICITY: Highly expressed in lung.  
CC -!- SIMILARITY: Belongs to the uteroglobin family. UGRP subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF274959; AAL25708.1; -;  
CC EMBL; AF274960; AAL25709.1; -;  
CC EMBL; AF274961; AAL25710.1; -;  
CC MGD; MGI:2153470; Scgb3a2.  
CC GO; GO:0005576; C:extracellular; IDA.  
CC GO; GO:0005576; F:protein binding; IPI.  
CC InterPro; IPR006038; Uteroglobin\_supf.  
CC Pfam; PF01093; Uteroglobin; 1.  
KW SIGNAL; Alternative splicing.  
FT SIGNAL 1 21 POTENTIAL.  
FT CHAIN 22 139 UTEROGLOBIN-RELATED PROTEIN 1.  
FT VARSPLIC 107 139 VSVLFLPMICAYPRDSKKQTAFIERVFSQKL -> EALS  
FT FTId=VSP\_006726.  
FT FTId=VSP\_006727.  
FT FTId=VSP\_006728.  
FT VARSPLIC 85 91 HLVS -> EALSHLV (in isoform A).  
FT VARSPLIC 92 139 Missing (in isoform A).  
FT FTId=VSP\_006728.  
SQ SEQUENCE 139 AA; 15431 MW; 8A2FB080B41E65E4 CRC64;  
Query Match 26.9%; Score 135; DB 1; Length 139;  
Best Local Similarity 35.1%; Pred. No. 2e-05;  
Matches 34; Conservative 17; Mismatches 36; Indels 10; Gaps 1;  
QY 1 MKLAALGLCVALS-  
Db 1 MKLVSIFFLLVTIGICGSATALLINRLPVDKLPV-----PLDDIIFSDPLKML 50  
QY 61 LSLIGIPVNHLEGSQKVAELGPQAVGAVKALKALL 97



```

Db 51 LKTLGIVSEHLVTGLKCKVDELGPSEASAVKLLVII 87
RESULT 5
ID TNK1_HUMAN STANDARD; PRT; 1327 AA.
AC Q95271; Q95272;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tankyrase 1 (EC 2.4.2.30) (TANK1) (Tankyrase I) (TNKS-1) (TRF1-
DE interacting ankyrin-related ADP-ribose polymerase).
GN TNKS OR TNKS1 OR TINK1 OR TINF1 OR PARP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RN [2]
RN [3]
RP MEDLINE=99454782; PubMed=10523501;
RA Smith S., de Lange T.;
RT "Cell cycle dependent localization of the telomeric PARP, tankyrase,
RT to nuclear pore complexes and centrosomes.";
RL J. Cell Sci. 112:3649-3656 (1999).
RN [3]
RP SUBCELLULAR LOCATION.
RP MEDLINE=99454782; PubMed=10523501;
RA Smith S., de Lange T.;
RT "Cell cycle dependent localization of the telomeric PARP, tankyrase,
RT to nuclear pore complexes and centrosomes.";
RL J. Cell Sci. 112:3649-3656 (1999).
RN [3]
RP FUNCTION, AND PHOSPHORYLATION.
RP MEDLINE=20556282; PubMed=10988299;
RA Chi N.-W., Lodish H.F.;
RT "Tankyrase is a Golgi-associated mitogen-activated protein kinase
RT substrate that interacts with IRAP in GLUT4 vesicles.";
RL J. Biol. Chem. 275:38437-38444 (2000).
RN [4]
RP FUNCTION, AND MUTAGENESIS OF HIS-1184 AND GLU-1291.
RP MEDLINE=21602874; PubMed=11739745;
RA Cook B.D., Dynek J.N., Chang W., Shostak G., Smith S.;
RT "Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2
RT at human telomeres.";
RL Mol. Cell. Biol. 22:332-342 (2002).
CC -!- FUNCTION: May regulate vesicle trafficking and modulate the
CC subcellular distribution of SLC2A4/GLUT4-vesicles. Has PARP
CC activity and can modify TRF1, and thereby contribute to the
CC regulation of telomere length.
CC -!- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-riboseyl}(N)-acceptor =
CC nicotinamide + {ADP-D-riboseyl}(N+1)-acceptor.
CC -!- SUBUNIT: Oligomerizes and associates with TNKS2. Interacts with
CC the cytoplasmic domain of INPEP/Otase in SLC2A4/GLUT4-vesicles.
CC binds to the N-terminus of telomeric TRF1 via the ANK repeats.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic, associated with the Golgi and
CC with juxtanuclear SLC2A4/GLUT4-vesicles. A minor proportion is
CC also found at nuclear pore complexes and around the pericentriolar
CC matrix of mitotic centrosomes. During interphase, a small fraction
CC of TNKS is found in the nucleus, associated with TRF1.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q95271-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q95271-2; Sequence=VSP_004538, VSP_004539;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Ubiquitous; highest levels in testis.
CC -!- PTM: Upon insulin-stimulation, phosphorylated on serine residues
CC by MAPK kinases.
CC -!- PTM: ADP-ribosylated (-auto).
CC -!- SIMILARITY: Belongs to the PARP family.
CC -!- SIMILARITY: Contains 15 ANK repeats.

```

```

CC -!- SIMILARITY: Contains 1 sterile alpha motif (SAM) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; AF082556; AAC79841.1; -
CC EMBL; AF082557; AAC79842.1; -
CC EMBL; AF082558; AAC79843.1; -
CC EMBL; AF082559; AAC79844.1; -
CC HSP; Q00420; IAWC.
CC Genew; HGNC:11941; TNKS.
CC MIM; 603303; -
CC GO; GO:0000781; C:Chromosome, telomeric region; IDA.
CC GO; GO:0003950; F:NAD ADP-ribosyltransferase activity; IDA.
CC GO; GO:000515; F:protein binding; IPI.
CC GO; GO:0007004; P:telomerase-dependent telomere maintenance; NAS.
CC InterPro; IPR002110; ANK.
CC InterPro; IPR001660; SAM.
CC Pfam; PF00023; ank; 19.
CC Pfam; PF00536; SAM; 1.
CC PRINTS; PR01415; ANKYRIN.
CC SMART; SM00248; ANK; 17.
CC SMART; SM00454; SAM; 1.
CC PROSITE; PS00088; ANK REPEAT; 15.
CC PROSITE; PS0297; ANK REP REGION; 1.
CC PROSITE; PS0105; SAM DOMAIN; 1.
CC Transferase; Glycosyltransferase; NAD; Golgi stack; Telomere;
CC Nuclear protein; Repeat; ANK repeat; ADP-ribosylation;
CC Phosphorylation; Alternative splicing.
CC REPEAT 215 247 ANK 1.
CC REPEAT 248 280 ANK 2.
CC REPEAT 281 313 ANK 3.
CC REPEAT 368 400 ANK 4.
CC REPEAT 401 432 ANK 5.
CC REPEAT 434 466 ANK 6.
CC REPEAT 521 556 ANK 7.
CC REPEAT 557 589 ANK 8.
CC REPEAT 590 622 ANK 9.
CC REPEAT 683 715 ANK 10.
CC REPEAT 716 748 ANK 11.
CC REPEAT 749 781 ANK 12.
CC REPEAT 836 868 ANK 13.
CC REPEAT 869 901 ANK 14.
CC REPEAT 902 934 ANK 15.
CC DOMAIN 1030 1059 SAM.
CC DOMAIN 1176 1327 PARP.
CC DOMAIN 9 14 POLY-HIS.
CC DOMAIN 27 34 POLY-PRO.
CC DOMAIN 128 134 POLY-SER.
CC DOMAIN 137 145 POLY-SER.
CC VARSPLIC 641 643 EST -> GHS (in isoform 2).
CC VARSPLIC 644 1327 /FTid=VSP_004538.
CC VARSPLIC 644 1327 Missing (in isoform 2).
CC MUTAGEN 1184 1184 H->A: LOSS OF ACTIVITY; WHEN ASSOCIATED
CC WITH A-1291.
CC MUTAGEN 1291 1291 E->A: LOSS OF ACTIVITY; WHEN ASSOCIATED
CC WITH A-1184.
CC SEQUENCE 1327 AA; 142010 MW; E14DE995C710B957 CRC64;
Query Match 16.5%; Score 83; DB 1; Length 1327;
Best Local Similarity 30.5%; Pred. No. 6.5;
Matches 29; Conservative 13; Mismatches 43; Indels 10; Gaps 3;
QY 15 CSSAAFLVGSAPVAPQVALESA-----EAGAGTLANPLGLTLPKLLISGLIPVN 69
DB 90 CTTTCTCTVAPVTV-PAVSTSAAGVAPNAGSGNNPSSSSSFTSSSSSSPSSPS 148

```

QY 70 HLIEGQ-----KCVAEIGPQAVGAVKALLGAL 100  
 Db 149 SLAESPEAGVSSATPLGGAAGPCTGVPVSGAL 183

## RESULT 6

TRD2\_STRCO  
 ID TRD2\_STRCO STANDARD; PRT; 335 AA.  
 AC Q9Z4W9;

DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Anthranilate phosphoribosyltransferase 2 (EC 2.4.2.18).  
 GN TRPD2 OR SCO3212 OR SCE8.05C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]

SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RX MEDLINE=21996410; PubMed=12000953;  
 RA Bentley S.D., Chater K.F., Cerrido-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieser T., Larke J., Murphy L., Oliver K., O'Neil S.,  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S., Taylor K.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares R., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,  
 RA Hopwood D.A.;

RT "Complete genome sequence of the model actinomycete Streptomyces

coelicolor A3(2).";

RL Nature 417:141-147(2002).

CC -! CATALYTIC ACTIVITY: Anthranilate + phosphoribosyldiphosphate =

CC N-5'-phosphoribosyl-anthranilate + diphosphate.

CC -! PATHWAY: Tryptophan biosynthesis; second step.

CC -! SIMILARITY: Belongs to the anthranilate phosphoribosyltransferase

CC family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AL939115; CAB38583.1; -.  
 CC PIR; T36304; T36304.

CC HAMAP; MF\_00211; -; 1.

CC InterPro; IPR005940; Ant phaspho trans.

CC InterPro; IPR000312; Glyco trans\_3.

CC Pfam; PF02885; Glycos trans\_3N; 1.

CC Pfam; PF00591; Glycos\_trans\_3; 1.

CC ProDom; PD001864; Glyco\_trans\_3; 1.

CC TIGRFAMs; TIGR01245; trpD; 1.

CC Tryptophan biosynthesis; Transferase; Glycosyltransferase;

CC Complete proteome.

CC SEQUENCE 335 AA; 34536 MW; 2706194E400B2F0D CRC64;

Query Match 15.3%; Score 77; DB 1; Length 335;  
 Best Local Similarity 34.5%; Pred. No. 6.5;  
 Matches 30; Conservative 9; Mismatches 36; Indels 12; Gaps 4;

QY 7 LGICVALSCSSAAAFVLSGA-----KPVAPQVAALSA--EAGACTLANPLGTL-NP 56  
 Db 117 LGVRIDLGEAAACLDRTGTFPLPAPVHP--AFRHTAGPRELGANTVFNLLGPLCNP 174

QY 57 LKLLLSLGIPIVNHLEIGSKQCVAEIG 83

Db 175 SGARLRTLGVPSELPVEPMTVEILRG 201

## RESULT 7

COPA\_HELPFE

ID COPA\_HELPFE STANDARD; PRT; 732 AA.

AC O32619;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Copper-transporting ATPase (EC 3.6.3.4).

GN COPA.

OS Helicobacter felis.

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;

OC Helicobacteraceae; Helicobacter.

OX NCBI\_TaxID=214;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=ATCC 49179;

RX MEDLINE=98101471; PubMed=9440521;

RA Bayle D., Waengler S., Weitzenegger T., Steinhilber W., Volz J.,

RA Przybylski M., Schaefer K.P., Sachs G., Melchers K.,

RT "Properties of the P-type ATPases encoded by the copA operons of

RT Helicobacter pylori and Helicobacter felis.";

RL J. Bacteriol. 180:317-329(1998).

CC -! FUNCTION: PROBABLY INVOLVED IN COPPER EXPORT.

CC -! CATALYTIC ACTIVITY: ATP + H(2)O + Cu(2+) (in) = ADP + phosphate +

CC Cu(2+) (out).

CC -! SUBCELLULAR LOCATION: Integral membrane protein.

CC -! SIMILARITY: Belongs to the cation transport ATPases family (P-type

CC ATPases) Subfamily IB.

CC -! SIMILARITY: Contains 1 HMA domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AJ001932; CAA05104.1; -.

CC PIR; T47269; T47269.

CC InterPro; IPR006403; ATPase-IB1 Cu.

CC InterPro; IPR006416; ATPase-IB Hvy.

CC InterPro; IPR001757; ATPase\_E1-E2.

CC InterPro; IPR001756; Cu ATPase.

CC InterPro; IPR008250; E1-E2 ATPase\_reg.

CC InterPro; IPR006121; HeavyMe transpt.

CC InterPro; IPR005834; Hydrolase.

CC Pfam; PF00122; E1-E2 ATPase; 1.

CC Pfam; PF00403; HMA; 1.

CC Pfam; PF00702; Hydrolase; 1.

CC PRINTS; PRO0119; CATATPASE.

CC PRINTS; PRO0943; CUATPASE.

CC TIGRFAMs; TIGR01511; ATPase-IB1 Cu; 1.

CC TIGRFAMs; TIGR01525; ATPase-IB Hvy; 1.

CC TIGRFAMs; TIGR01494; ATPase P-type; 3.

CC PROSITE; PS00154; ATPASE\_E1-E2; 1.

CC PROSITE; PS01047; HMA\_1; 1.

CC PROSITE; PS0846; HMA\_2; 1.

CC Transprot; Ion transport; Copper; Magnesium; Phosphorylation.

CC Metal-binding; Copper; Magnesium; Phosphorylation.

CC DOMAIN 1 88

FT TRANSMEM 89 109

FT DOMAIN 110 122

FT TRANSMEM 123 142

FT DOMAIN 143 149

FT TRANSMEM 150 170

FT DOMAIN 171 187

FT TRANSMEM 188 208

FT DOMAIN 209 336

FT TRANSMEM 337 359

FT DOMAIN 360 365

FT EXTRACELLULAR (POTENTIAL).

FT POTENTIAL.

FT CYTOPLASMIC (POTENTIAL).

FT EXTRACELLULAR (POTENTIAL).

FT POTENTIAL.

FT CYTOPLASMIC (POTENTIAL).

FT POTENTIAL.

FT EXTRACELLULAR (POTENTIAL).

```

FT TRANSMEM 366 383 POTENTIAL.
FT DOMAIN 384 663 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 664 683 POTENTIAL.
FT TRANSMEM 684 694 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 695 713 POTENTIAL.
FT DOMAIN 714 732 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 732 69 HNA.
FT MOD RES 421 421 PHOSPHORYLATION (BY SIMILARITY).
FT METAL 13 13 COPPER (POTENTIAL).
FT METAL 16 16 COPPER (POTENTIAL).
FT METAL 609 609 MAGNESIUM (BY SIMILARITY).
FT METAL 613 613 MAGNESIUM (BY SIMILARITY).
FT METAL 732 613 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 732 AA; 78853 MW; 7105107EA943EFD CRC64;

Query Match 15.1%; Score 76; DB 1; Length 732;
Best Local Similarity 26.3%; Pred. No. 16;
Matches 35; Conservative 17; Mismatches 39; Indels 42; Gaps 6;

QY 6 LLGLCVLSCSSAAFLVGSAPVQVAAAL---ESAAEAGAG-----45
DB 447 LLTLCASLEAQSEHVIAKGIVAHAKGEGTALQEVQAKPGFGIKGVVGDOIKAAGNLE 506
QY 46 --TIANPLGTNPKLLSS-----LGIPV--NHLIEGSKCVAEILGPQAVGA-----89
DB 507 FFNLNPFGLTEGIVQVGTETQILGVWLADSLKEGSEKAISEL--KALGVKTKLLSSD 564
QY 90 ----VKALKALLG 98
DB 565 NLENVRALATOLG 577

RESULT 8
MDC1_HUMAN
ID MDC1_HUMAN STANDARD; PRT; 362 AA.
AC Q9H1K6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mesoderm development candidate 1.
GN MESDC1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Capleton M., Soares M.B., Donald M.P., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustin T.B., Roehyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smallos D.E.,
Schnierch A., Schein J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: SOME, TO TALIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC -----
CC EMBL; AY007810; AAG41058.1; -
CC Genew; HGNC:13519; MESDC1.
SQ SEQUENCE 362 AA; 37758 MW; 37BF391D663B1D6E CRC64;

Query Match 14.9%; Score 75; DB 1; Length 362;
Best Local Similarity 35.4%; Pred. No. 11;
Matches 28; Conservative 6; Mismatches 33; Indels 12; Gaps 3;

QY 3 LAALLGLCVLSCSSAAFLVGSAPVQVAAAL---VAALSAAGAGATL-ANP 50

```

```

DB 112 LVLEGLDVVSLTECSAAHYLAAVATPAQPAQGLVDRVYVTRCRHEVGGCAVLRATP 171
QY 51 LGTLNPLKLLSSLGIPVN 69
DB 172 LADMTFOLLLEVSQGLSRN 190

RESULT 9
MDC1_MOUSE
ID MDC1_MOUSE STANDARD; PRT; 362 AA.
AC Q9ERE8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Mesoderm development candidate 1.
GN MESDC1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RP STRAIN=129/SvJ;
RX MEDLINE=21145589; PubMed=11247670;
Wines M.E., Lee L., Katari M.S., Zhang L., DeRossi C., Shi Y.,
Perkins S., Feldman M., McCombie W.R., Holdener B.C.;
"Identification of mesoderm development (mesd) candidate genes by
comparative mapping and genome sequence analysis."
RL Genomics 72:88-98(2001).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Capleton M., Soares M.B., Donald M.P., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Ustin T.B., Roehyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smallos D.E.,
Schnierch A., Schein J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: SOME, TO TALIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC -----
CC EMBL; AF311213; AAG33620.1; -
CC Genew; BC018326; AAH18326.1; -
CC MGDI; MGDI:1891420; Mesdcl.
SQ SEQUENCE 362 AA; 37786 MW; EFE9BBFC09BB7CB5 CRC64;

Query Match 14.9%; Score 75; DB 1; Length 362;
Best Local Similarity 35.4%; Pred. No. 11;
Matches 28; Conservative 6; Mismatches 33; Indels 12; Gaps 3;

QY 3 LAALLGLCVLSCSSAAFLVGSAPVQVAAAL---VAALSAAGAGATL-ANP 50

```

```
Db 112 LVELGDLVVSIECSAHAYLAATVATGQAQPGVLVDVRYVTRCRHEVEQGCAVLRAATP 171
QY 51 LGTLNPLKLLSSLGIPVN 69
Db 172 LADMTFQLLEVSQGLSRN 190

RESULT 10
SNIL_MOUSE
ID SNIL_MOUSE STANDARD; PRT; 779 AA.
AC Q60670;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Probable serine/threonine protein kinase SNF1LK (EC 2.7.1.1) (HRT-20)
DE (Myocardial SNF1-like kinase).
GN SNF1LK OR MSK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND REVISIONS TO 8; 16-18; 44-45; 316 AND 435.
RC TISSUE=Embryo;
RA Ruiz J.C.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-435 FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=Embryo;
RX MEDLINE=95200798; PubMed=7893599;
RA Ruiz J.C., Conlon F.L., Robertson E.J.;
RT "Identification of novel protein kinases expressed in the myocardium
of the developing mouse heart.";
RL Mech. Dev. 48:153-164(1994).
CC -!- TISSUE SPECIFICITY: Expressed in lung, skin, ovary, heart and
CC stomach. No expression in brain, liver or skeletal muscle.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNF1
CC subfamily.
CC -!- SIMILARITY: Contains 1 UBA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U11494; AAA67926.2; -
CC HSP; P24941; IAKL.
CC MGD; MG1:104754; Snflk.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_Thr_kinase_AS.
CC InterPro; IPR002290; Ser_thr_kinase.
CC InterPro; IPR001245; Tyr_kinase.
CC InterPro; IPR000449; UBA domain.
CC Pfam; PF00669; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS50030; UBA; 1.
CC Transferase; Serine/threonine-protein kinase; ATP-binding.
KW DOMAIN 27 278 PROTEIN KINASE.
FT DOMAIN 303 343 UBA.
FT NP_BIND 33 41 ATP [BY SIMILARITY].
FT BINDING 56 56 ATP [BY SIMILARITY].
FT ACT_SITE 149 149 BY SIMILARITY.
FT ACT_SITE 149 149 BY SIMILARITY.
SQ SEQUENCE 779 AA; 85027 MW; 7B08131BC46D9C4E CRC64;
```

```
Query Match 14.7%; Score 74; DB 1; Length 779;
Best Local Similarity 32.9%; Pred. No. 26;
Matches 26; Conservative 11; Mismatches 30; Indels 12; Gaps 3;

QY 10 CVALSCSSAAFLVGSAPVQAQVAALESAAEAGATLANP--LGTNPKLLLSLGP 67
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 490 CIIVSSGATASPSSECTSGSDCLPFSASEGPAGLGSG-LATPGLLTSSPVRLASPEL--- 545
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 68 VNHLEGSQKCVAEIGPOA 86
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 546 -----GQSATPVLQQA 558
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
DXS_WIGBR
ID DXS_WIGBR STANDARD; PRT; 626 AA.
AC Q8D557;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 1-deoxy-D-xylulose 5-phosphate synthase (EC 2.2.1.7) (1-
DE deoxyxylulose-5-phosphate synthase) (DXP synthase) (DXPS).
GN DXS OR WIGBR1440
OS Wigglesworthia glossinidia brevipalpis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Wigglesworthia.
OX NCBI_TaxID=36870;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22297718; PubMed=12219091;
RA Akman L., Yamashita A., Watanabe H., Oshima K., Shiba T., Hattori M.,
RA Aksoy S.;
RT "Genome sequence of the endocellular obligate symbiont of tsetse
RT flies, Wigglesworthia glossinidia.";
RL Nat. Genet. 32:402-407(2002).
CC -!- FUNCTION: Catalyzes the acyloln condensation reaction between C
CC atoms 2 and 3 of pyruvate and glyceraldehyde 3-phosphate to yield
CC 1-deoxy-D-xylulose-5-phosphate (DXP) (By similarity).
CC -!- CATALYTIC ACTIVITY: Pyruvate + D-glyceraldehyde 3-phosphate = 1-
CC deoxy-D-xylulose 5-phosphate + CO(2).
CC -!- COFACTOR: Binds 1 thiamine pyrophosphate per subunit (By
CC similarity).
CC -!- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; first step.
CC -!- PATHWAY: Biosynthetic pathway to thiamine and pyridoxol; first
CC step.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SIMILARITY: Belongs to the transketolase family. DXPS subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB063521; BAC24290.1; -
CC HAMAP; MF_00315; - 1.
CC InterPro; IPR005477; Dxs_synth.
CC InterPro; IPR009014; Transketol_C_like.
CC InterPro; IPR005476; Transketolase_C.
CC InterPro; IPR005475; Transketolase_CR.
CC InterPro; IPR005474; Transketolase_N.
CC Pfam; PF02779; transket_pvr; 1.
CC Pfam; PF02780; transketolase_C; 1.
CC TIGRfam; TIGR00204; dxe; 1.
CC PROSITE; PS00802; TRANSKETOLASE_1; 1.
CC PROSITE; PS00802; TRANSKETOLASE_2; FALSE_NEG.
KW Transferase; Flavoprotein; Thiamine pyrophosphate;
KW Isoprene biosynthesis; Thiamine biosynthesis; Complete proteome.
SQ SEQUENCE 626 AA; 69851 MW; C77008D282ABE21 CRC64;
```

Query Match 14.6%; Score 73.5; DB 1; Length 626;

```
Best Local Similarity 30.0%; Pred. No. 24;
Matches 24; Conservative 14; Mismatches 41; Indels 1; Gaps 1;

QY 22 LVSAKPVQVVAALSAEAGAGTIANPLGTLNPKLLSLGIPVNHLEGSQKCV-A 80
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 546 LIRMAKTHKALITTEENVINGGAGSAVNEFIMYNKLLVPLNIGIPDNFVSHGSGTEVRS 605

QY 81 ELGPQAVGAVKALKALGAL 100
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 606 SLGLDSIGIKKINWLNFL 625

RESULT 12
SN1L_RAT STANDARD; PRT; 776 AA.
AC QRIJ5; Q9R081;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable serine/threonine protein kinase SNF1LK (EC 2.7.1.-) (Salt-
DE inducible protein kinase) (Protein kinase KID2).
GN SNF1LK OR SIK OR KID2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
STRAIN=Sprague-Dawley; TISSUE=Adrenal gland;
MEDLINE=9330184; PubMed=10403390;
PA Wang Z., Takemori H., Haider S.K., Nonaka Y., Okamoto M.;
RT "Cloning of a novel kinase (SIK) of the SNF1/AMPK family from high
RT salt diet-treated rat adrenal.";
RL FEBS Lett. 453:135-139 (1999).
[2]
SEQUENCE FROM N.A.
PA Feldman J.D., Vician L., Crispino M., Hoe W., Baudry M.,
PA Herschman H.R.;
RT "The KID2 gene encodes a protein kinase induced by depolarization in
RT brain.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. SNF1
CC subfamily.
CC -!- SIMILARITY: Contains 1 UBA domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB020480; BAA92673.1; -
CC EMBL; AF106937; AAF14191.1; -
CC HSSP; P24941; 1AQ1.
CC InterPro; IPR000719; Prot kinase.
CC InterPro; IPR008271; Ser Thr pkin AS.
CC InterPro; IPR002290; Ser Thr pkinase.
CC InterPro; IPR000449; UBA_domain.
CC Pfam; PF00063; Pkinase; 1.
CC ProDom; PD000001; Prot kinase; 1.
CC SMART; SMO0220; S_TKC; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS50030; UBA; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 27 278 PROTEIN_KINASE.
FT DOMAIN 303 343 UBA.
FT NP_BIND 33 41 ATP (BY SIMILARITY).
FT BINDING 56 56 ATP (BY SIMILARITY).
FT ACT_SITE 149 149 BY SIMILARITY.
FT
```

```
FT CONFLICT 473 473 R -> K (IN REF. 2).
SQ SEQUENCE 776 AA; 84908 MW; 7BF745AF28F17E6E CRC64;

Query Match 14.6%; Score 73.5; DB 1; Length 776;
Best Local Similarity 33.7%; Pred. No. 28;
Matches 28; Conservative 11; Mismatches 31; Indels 13; Gaps 4;

QY 10 CVALSCSSAAFLVGSAXPVQVVAALSAEAGAGTIANP--LGTNPLKLLSLGIP 67
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 490 CIIVSSAAVSPSEGTSSDCLFFSASGPAAGLGG--LATPGLLGTSFVRLASPFPL--- 545

QY 68 VNHLEGSQKCVAEELGQA-VGA 89
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 546 -----GQSATPVLQSQAGLGA 562

RESULT 13
SYA_THETH STANDARD; PRT; 882 AA.
ID P4941;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (ALARS).
GN ALAS.
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OX NCBI_TaxID=274;
RN [1]
SEQUENCE FROM N.A.
STRAIN=HB8 / ATCC 27634;
MEDLINE=97351141; PubMed=9207019;
PA Lechner A., Martin A., Zuleeg T., Limmer S., Kreutzer R.;
RA "A biologically active 53 kDa fragment of overproduced alanyl-tRNA
RT synthetase from Thermus thermophilus HB8 specifically interacts with
RT tRNA ala acceptor helix.";
RL Nucleic Acids Res. 25:2737-2744 (1997).
CC -!- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
CC diphosphate + L-alanyl-tRNA(Ala).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y08363; CAA69650.1; -
CC HAMAP; MF_00036; -; 1.
CC InterPro; IPR003156; DHHA1.
CC InterPro; IPR002318; tRNA-synt_2C.
CC InterPro; IPR006193; tRNA_synt_Ala.
CC Pfam; PF02272; DHHA1; 1.
CC Pfam; PF01411; tRNA-synt_2C; 1.
CC PRINTS; P000980; TRNASYNTHALA.
CC TIGRfams; TIGR00344; alaS; 1.
CC PROSITE; PS50860; AA tRNA LIGASE II ALA; 1.
CC Aminoacyl-tRNA synthetase; Protein Biosynthesis; Ligase; ATP-binding.
KW Aminoacyl-tRNA synthetase; Protein Biosynthesis; Ligase; ATP-binding.
SQ SEQUENCE 882 AA; 97454 MW; E60E3E34480CFC56 CRC64;

Query Match 14.6%; Score 73.5; DB 1; Length 882;
Best Local Similarity 32.6%; Pred. No. 32;
Matches 30; Conservative 12; Mismatches 47; Indels 3; Gaps 2;

QY 6 LLGLCVALSASSAAFLVGSAXPVQVVAALSAEAGAGTIANPLGTLNPKLLSLG 65
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 681 LCGGCHVRRGTGTCALFIRSEAVSAGVRRIR--AVTGEAIRFARGSLNRKLAERLE 738

QY 66 IPVNHLEGSQKCVAEELGQA-VGA 97
```



```
FT DOMAIN 139 153 POLY-HIS.
FT DOMAIN 156 162 POLY-HIS.
FT DOMAIN 163 172 POLY-PRO.
FT DOMAIN 217 231 POLY-ALA.
FT DOMAIN 249 258 POLY-ALA.
FT DOMAIN 321 330 POLY-ALA.
FT DOMAIN 396 399 POLY-ALA.
SQ SEQUENCE 428 AA; 45170 MW; DB8A8EFD1B94AB10 CRC64;

Query Match 14.1%; Score 71; DB 1; Length 428;
Best Local Similarity 29.2%; Pred. No. 28;
Matches 33; Conservative 17; Mismatches 43; Indels 20; Gaps 6;

QY 11 VALSCSAAFLVGSAPVAQ-PVALESAAAGAGTLANPLGTINPL-----KL 59
Db 220 VAAAAAAAAAAGVSGRLSQFPYGLGSAAAAAAAAAASTTGFKHPAENIIGRDYKG 279
QY 60 LLSLSLGIPI-----VNHL---IEGS-QKCVIELGPOAVGAKALKALLGALTVEG 104
Db 280 VLQAGGLPLASVHHLGYPFVGQLSNVGVSMVPH-VGVNDSVAAAAAANAAG 331

RESULT 16
DNAME RHIME STANDARD; PRT; 480 AA.
AC P35890;
DT 01-JUN-1994 (Rel. 29, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chromosomal replication initiator protein dnaA.
GN DAAA OR R00368 OR SMC01167.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -!- FUNCTION: Plays an important role in the initiation and regulation
CC of chromosomal replication. Binds to the origin of replication; it
CC binds specifically double-stranded DNA at a 9 bp consensus (dnaA
CC box): 5'-TTATC(C/A)(C/A)A-3'. DnaA binds to ATP and to acidic
CC phospholipids.
CC -!- SIMILARITY: Belongs to the dnaA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L25439; AAA26258.1; ALT_INIT.
DR EMBL; L39265; AAA91097.1; ALT_INIT.
DR EMBL; AL591783; CAC41805.1; ALT_INIT.
```

```
DR HAMAP; MF 00377; -, 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001957; Bac DnaA.
DR Pfam; PF00308; bac_dnaA; 1.
DR PRINTS; PR00051; DNaA.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00362; DnaA; 1.
DR PROSITE; PS01008; DNaA; 1.
KW DnaA replication; DNA-binding; ATP-binding; Complete proteome.
FT NP BIND 181 188 ATP (POTENTIAL).
SQ SEQUENCE 480 AA; 53579 MW; A346219949FAD6A3 CRC64;

Query Match 14.1%; Score 71; DB 1; Length 480;
Best Local Similarity 38.9%; Pred. No. 31;
Matches 21; Conservative 11; Mismatches 14; Indels 8; Gaps 3;

QY 20 AFLVGSAPVAQFVALESAAAGAGTLANPLGTINPLKLLSLGIPVNHLE 73
Db 147 SFVEGSSNVA--LAAARTIAEAGAGAV-----RPNPL-FIHSSVGLGKTHLLQ 192

RESULT 17
RL7_RALSO
ID RL7_RALSO STANDARD; PRT; 124 AA.
AC Q8KUZ7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L7/L12.
GN RPLL OR RSC3035 OR RSC4722.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GMI1000;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.-C., Cattolico L.,
RA Chandler M., Choine N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
CC -!- FUNCTION: Seems to be the binding site for several of the factors
CC involved in protein synthesis and appears to be essential for
CC accurate translation (By similarity).
CC -!- SIMILARITY: Belongs to the L12P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL646073; CAD16744.1; -.
DR HAMAP; MF 00368; -, 1.
DR InterPro; IPR0008932; Ribos_L12/7_olig.
DR InterPro; IPR002026; Ribosomal_L12.
DR Pfam; PF00542; Ribosomal_L12; 1.
DR ProDom; PD001326; Ribosomal_L12; 1.
DR TIGRFAMs; TIGR00855; L12; 1.
DR Ribosomal protein; Complete proteome.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 124 AA; 14520 MW; 271C3C71F80269E CRC64;

Query Match 13.7%; Score 69; DB 1; Length 124;
Best Local Similarity 31.8%; Pred. No. 14;
Matches 28; Conservative 17; Mismatches 35; Indels 8; Gaps 5;
```







```

DR EMBL; X57471; CAA40707.1; -.
DR EMBL; X59402; CAA42043.1; -.
DR EMBL; AE000164; AAC73691.1; -.
DR EMBL; U82598; AAB40789.1; ALT_INIT.
DR PIR; S16296; S16296.
DR EcoGene; EGI0286; fepD.
DR InterPro; IPR000515; BPD_transp.
DR InterPro; IPR000522; FecD.
DR Pfam; PF01032; FecCD; 1.
DR ProDom; PD001557; FecCD; 1.
KW Iron transport; Transpote; Inner membrane; Transmembrane;
KW Complete proteome.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 64 84 POTENTIAL.
FT TRANSMEM 93 113 POTENTIAL.
FT TRANSMEM 121 141 POTENTIAL.
FT TRANSMEM 152 172 POTENTIAL.
FT TRANSMEM 193 213 POTENTIAL.
FT TRANSMEM 242 262 POTENTIAL.
FT TRANSMEM 280 300 POTENTIAL.
FT TRANSMEM 306 326 POTENTIAL.
FT TRANSMEM 334 354 POTENTIAL.
FT CONFLICT 233 234 RV -> AL (IN RFP. 2).
SQ SEQUENCE 334 AA; 33871 MW; BB5C5E939A96951B CRC64;

Query Match 13.5%; Score 68; DB 1; Length 334;
Best Local Similarity 22.1%; Pred. No. 42;
Matches 34; Conservative 19; Mismatches 43; Indels 58; Gaps 6;

QY 3 LAALLGLCVALS-----SAAFLVGS-----KPAQP----- 32
DB 41 LEAFSGTCSADCTIVLDARLPRTLGLAGGALGLAGALMOTLRNPLADPGLGVNAG 100
QY 33 -----VAALSAABAGAGTLANPL-----GTLNPKLLSLGIPV 68
DB 101 ASPAIVGAALFGYSRAQQLNAPAGALVASLIVFTSGSQGGQLSPVRLTLA--GVAL 158
QY 69 NHIEGSKVCVSELGQAVGAVKALGALTU 102
DB 159 AAVLEGLTSGIALNPDVYDQLAFWQA--GSLDI 190

RESULT 20
SILF MOUSE
ID -SILF MOUSE STANDARD; PRT; 569 AA.
AC Q920G3;
DC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sialic acid binding Ig-like lectin-F precursor (mSilec-F).
GN SIGLEC-F.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=21576254; PubMed=11579105;
RA Argata T., Hingorani R., Varki N.M., Varki A.;
RT "Cloning and characterization of a novel mouse Siglec-F:
RT differential evolution of the mouse and human (CD33) Siglec-3-related
RT gene clusters."
EL J. Biol. Chem. 276:45128-45136 (2001).
CC -!- FUNCTION: Putative adhesion molecule that mediates sialic-acid
CC dependent binding to cells. Preferentially binds to alpha2,3-
CC linked sialic acid. The sialic acid recognition site may be masked
CC by cis interactions with sialic acids on the same cell surface.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Predominantly expressed by immature
CC monocyte/myeloid lineage cells in bone marrow. Also found at
CC lower levels in mature neutrophils and monocytes.
CC -!- DOMAIN: Contains 1 copy of a cytoplasmic motif that is referred
CC to as the immunoreceptor tyrosine-based inhibitor motif (ITIM).

```

```

CC This motif is involved in downmodulation of cellular responses.
CC The phosphorylated ITIM motif binds to the SH2 domain of
CC PTPN6/SHP-1.
CC -!- SIMILARITY: Belongs to the immunoglobulin superfamily. SIGLEC
CC (sialic acid binding Ig-like lectin) family.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF293371; AAL11043.1; -.
CC InterPro; IPR007110; IG-like.
CC InterPro; IPR003598; IG_c2.
CC Pfam; PF00047; IG; 2.
CC SMART; SM00408; IGC2; 1.
CC PROSITE; PS50835; IG_LIKE; 2.
KW Cell adhesion; Lectin; Transmembrane; Signal; Glycoprotein;
KW Immunoglobulin domain; Repeat.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 569 SIALIC ACID BINDING IG-LIKE LECTIN-F.
FT DOMAIN 17 439 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 440 460 POTENTIAL.
FT DOMAIN 461 569 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 18 116 IG-LIKE V-TYPE.
FT DOMAIN 139 224 IG-LIKE C2-TYPE 1.
FT DOMAIN 229 324 IG-LIKE C2-TYPE 2.
FT SITE 536 541 ITIM MOTIF.
FT SITE 559 564 SLAM-LIKE MOTIF.
FT DISULFID 35 163 BY SIMILARITY.
FT DISULFID 40 96 BY SIMILARITY.
FT DISULFID 157 206 BY SIMILARITY.
FT DISULFID 265 308 BY SIMILARITY.
FT CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 203 203 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 369 369 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 372 372 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 569 AA; 61476 MW; 8093838090484FC1 CRC64;

Query Match 13.5%; Score 68; DB 1; Length 569;
Best Local Similarity 29.9%; Pred. No. 67;
Matches 26; Conservative 11; Mismatches 38; Indels 12; Gaps 3;

QY 3 LAALLGLCVALS-----SAAFLVGSAPVQPVAALESAAEAGAGTLANPLGTLNPKLLLS 62
DB 447 LMALLAVCLC-----LIPTTVKVLKKSAKVAATKGNHLAKNPASTINSASITSS 497
QY 63 --SLGIPV-NHLIEGSKVCVSELGPOA 86
DB 498 NIALGPIQGHLEPQSQKQKQPPPLA 524

RESULT 21
Y233 HUMAN
ID Y233 HUMAN STANDARD; PRT; 2035 AA.
AC Q92508;
DC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein KIAA0233.
GN KIAA0233.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

```

RA SEQUENCE FROM N.A.  
RP TISSUE=Bone marrow;  
RC MEDLINE=97191544; PubMed=9039502;  
RX Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayashi Y.,  
RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.,  
RT "Prediction of the coding sequences of unidentified human genes. VI.  
RT The coding sequences of 80 new genes (K1AA0201-K1AA0280) deduced by  
RT analysis of cDNA clones from cell line KG-1 and brain.",  
RL DNA Res. 3:321-329(1996).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D87071; BAAL3240.1; --  
KW Hypothetical protein.  
FT DOMAIN 260 264 POLY-GLN.  
FT DOMAIN 265 270 POLY-GLU.  
FT DOMAIN 450 453 POLY-LEU.  
FT DOMAIN 618 621 POLY-LEU.  
FT DOMAIN 940 944 POLY-GLU.  
FT DOMAIN 1409 1418 POLY-GLU.  
FT DOMAIN 1487 1490 POLY-ILE.  
FT DOMAIN 1680 1683 POLY-LYS.  
SQ SEQUENCE 2035 AA; 233040 MW; F337E333DFC0CB88 CRC64;  
Query Match 13.4%; Score 67.5; DB 1; Length 2035;  
Best Local Similarity 29.2%; Pred. No. 2.3e+02;  
Matches 19; Conservative 12; Mismatches 29; Indels 5; Gaps 2;  
QY 30 AQPVALESAAEA--GAGTLANPLGTLNPLKLLSSLGIPVN---HLIEGSKQCVAEELGP 84  
Db 1085 SQAEATLPQTEAPNAPATVSGSLGAEPLSSMTDDNGSLSTGYHTRSGSEEAVIDPGE 1144  
QY 85 QAVGA 89  
Db 1145 REAGA 1149  
RESULT 22  
GCP MYCLE STANDARD; PRT; 351 AA.  
ID GCP MYCLE  
AC P37969; Q49725; PRT; 351 AA.  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Probable O-sialoglycoprotein endopeptidase (EC 3.4.24.57)  
DE (Glycoprotease).  
GN GCP OR ML0379 OR U229E OR U1620C OR B229\_C3\_246 OR BL620\_C3\_226.  
OS Mycobacterium leprae.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1769;  
RN (1)  
RP SEQUENCE FROM N.A.  
RA Smith D.R., Robinson K.;  
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
RN (2)  
RP SEQUENCE FROM N.A.  
RC STRAIN=TN;  
RX MEDLINE=21128732; PubMed=11234002;  
RA Cole S., Eiglmier K., Parkhill J., James K.D., Thomson N.R.,  
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hanlin N.,  
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,  
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,

RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
RA Barrell B.G.;  
RT "Massive gene decay in the leprosy bacillus";  
RL Nature 403:1007-1011(2001).  
RN (3)  
RP SEQUENCE OF 286-351 FROM N.A.  
RX MEDLINE=92374850; PubMed=1354834;  
RA de Wit T.F.R., Bekelle S., Osland A., Miko T.L., Hermans P.W.M.,  
RA van Soelingen D., Drijfhout J., Schoenlingh R., Janson A.A.M.,  
RA Thole J.E.R.;  
RT "Mycobacteria contain two groEL genes: the second Mycobacterium  
RT leprae groEL gene is arranged in an operon with groES";  
RL Mol. Microbiol. 6:1995-2007(1992).  
CC -|- FUNCTION: Could be a metalloprotease.  
CC -|- CATALYTIC ACTIVITY: Hydrolysis of O-sialoglycoproteins; cleaves  
CC 31-Arg-Asp-32 bond in glycophorin A. Does not cleave  
CC unglycosylated proteins, desialylated glycoproteins or  
CC glycoproteins that are only N-glycosylated.  
CC -|- COFACTOR: Zinc (Probable).  
CC -|- SIMILARITY: Belongs to peptidase family M22.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U00020; AAA17310.1; ALT SEQ.  
DR EMBL; AL583918; CAC29887.1; --  
DR EMBL; U00015; AAC43226.1; --  
DR PIR; S72817; S72817.  
DR Leproma; ML0379; --  
DR MEROPS; M22.UFW; --  
DR InterPro; IPR000905; Peptidase\_M22.  
DR Pfam; PF00814; Peptidase\_M22; 1.  
DR PIRSF; PIRSF004537; OsiAlgC\_ptds; 1.  
DR PRINTS; PR00789; OSIALOPTASE  
DR ProDom; PD002367; Peptidase\_M22; 1.  
DR TIGRFAMs; TIGR00329; GCP; 1.  
DR PROSITE; PS01016; GLYCOPROTEASE; 1.  
KW Hydrolase; Metalloprotease; Zinc; Complete proteome.  
FT METAL 124 124 ZINC (POTENTIAL).  
FT METAL 128 128 ZINC (POTENTIAL).  
SQ SEQUENCE 351 AA; 35405 MW; F76A14B739A7BAE4 CRC64;  
Query Match 13.3%; Score 67; DB 1; Length 351;  
Best Local Similarity 34.5%; Pred. No. 53;  
Matches 29; Conservative 6; Mismatches 39; Indels 10; Gaps 2;  
QY 4 AALLGLCVLSCSSAAAFVGSAPVAPVAALESAAEAGAGTLANPLGTLNPLKLLSS 63  
Db 62 AHLALGPTTICALAAAGLTGSAKPDV-----VAATIGPLAGALLGVAAAKAYSAA 114  
QY 64 LGIP---VNHIEGSKQCVAEELGP 84  
Db 115 WGVPPFVAVNHIGHLAADVYEHGP 138  
RESULT 23  
FUNG\_PESN  
ID FUNG\_PESN STANDARD; PRT; 464 AA.  
AC Q885V0;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Fumarate hydratase class II (EC 4.2.1.2) (Fumarase C).  
GN FUMC-1 OR PSF01731.  
OS Pseudomonas syringae (pv. tomato).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=323;

[1]
SEQUENCE FROM N.A.
MEDLINE=22834015; PubMed=12928499;
RX STRAIN=DC3000;
RA Buell C.R., Joadar V., Lindeberg M., Selenog J., Paulsen I.T.,
RA Winn M.L., Dodson R.J., Deby R.T., Durkin A.S., Kolonay J.F.,
RA Madupu R., Daugherty S., Brinkac L., Beanan M.J., Haft D.H.,
RA Nelson W.C., Davidson N., Tran B., Russell D., Barry K., Utterback T.,
RA Khouri H., Fedorova N., Zafar N., Liu J., Yuan Q.,
RA Van Aken S.E., Feldblyum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,
RA Alfano J.R., Cartinhouer S., Chatterjee A.K., Delaney T.P.,
RA Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,
RA White O., Fraser C.M., Collier A.,
RA "The complete genome sequence of the Arabidopsis and tomato pathogen
RT Pseudomonas syringae pv. tomato DC3000.";
Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
CC -1- CATALYTIC ACTIVITY: (S)-malate + fumarate + H(2)O.
CC -1- PATHWAY: Tricarboxylic acid cycle.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the class-II fumarate / aspartase family.
Fumarase subfamily.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL: AE016862; AAC55251.1; -
TIGR: PSPT01731; -
HAWP: MF\_00743; -
InterPro: IPR000362; Fumarate\_lyase.
InterPro: IPR008948; L-Aspartase-like.
Pfam: PF00206; lyase 1; 1.
PRINTS: PR00149; FUMARATELYASE.
PROSITE: PS00163; FUMARATE LYASES; 1.
Lysase: Tricarboxylic acid cycle; Complete proteome.
ACT SITE 186 186 POTENTIAL
BINDING 322 322 SUBSTRATE CARBOXYL (POTENTIAL).
SEQUENCE 464 AA; 49077 MW; 6B12FF7257FFD2 CRC64;
Query Match 13.3%; Score 67; DB 1; Length 464;
Best Local Similarity 31.3%; Pred. No. 68;
Matches 31; Conservative 14; Mismatches 36; Indels 18; Gaps 5;
QY 2 KLAALLGLCVALSASSAAFLVGSAPVQVPALESAAEAGAGTLANPLGTINPLKLL- 60
DB 247 ELAALSLGPFVTPNPKFAA-LAGH-----EPLAALSGALKTLAGTL---MKIANDLRLIG 297
QY 61 -----LSSLGIPVNHLEGSQKCVBELGPQAVGAKAL 93
DB 298 SGPRAGLAELVLPANE--PGSSIMPGKVNPTCEALSML 334
RESULT 24
ID TA2R RAT STANDARD; PRT; 341 AA.
AC P34978;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Thromboxane A2 receptor (TXA2-R) (Prostanoid TP receptor) (TXR2).
GN TBXA2R.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI\_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
STRAIN=Sprague-Dawley; TISSUE=Kidney;
MEDLINE=95362824; PubMed=7635958;
RX Abe T., Takeuchi K., Takahashi N., Teuteumi E., Taniyama Y., Abe K.;
RT "rat kidney thromboxane receptor: molecular cloning, signal
transduction, and intrarenal expression localization.";
J. Clin. Invest. 96:657-664(1995).
RN [2]
SEQUENCE FROM N.A.
STRAIN=Sprague-Dawley; TISSUE=Brain;
MEDLINE=95210358; PubMed=7696353;
RX Kitamura J., Hashimoto H., Sugimoto Y., Sawada M., Negishi M.,
RA Kizumura A., Marunouchi T., Ichikawa A., Baba A.;
RT "cDNA cloning of a thromboxane A2 receptor from rat astrocytes.";
Biochim. Biophys. Acta 1265:220-223(1995).
RN [3]
SEQUENCE FROM N.A.
STRAIN=Sprague-Dawley; TISSUE=Kidney;
MEDLINE=97090677; PubMed=8936885;
RX D'Angelo D.B., Texasawa T., Carlisle S.J., Dorn G.W. II, Lynch K.R.;
RA "Characterization of a rat kidney thromboxane A2 receptor: high
affinity for the agonist ligand I-BOP.";
Prostaglandins 52:303-316(1996).
CC -1- FUNCTION: Receptor for thromboxane A2 (TXA2), a potent stimulator
of platelet aggregation. The activity of this receptor is mediated
by a G-protein that activate a phosphatidylinositol-calcium second
messenger system. In the kidney, the binding of TXA2 to glomerular
TP receptors causes intense vasoconstriction.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: In the brain, expressed in all types of glial
cells. In the kidney, expressed in the mesangial cells of the
glomerulus, smooth muscle cells of the renal arterioles, and in
transitional cell epithelium of renal pelvis.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL: D21158; BAA04694.1; -
EMBL: D32080; BAA06844.1; -
PIR: I55623; I55623.
InterPro: IPR000276; GPCR\_Rhodopsn.
InterPro: IPR008365; ProstanoidR.
Pfam: PF00001; 7tm 1; 1.
PRINTS: PR01788; PROSTANOIDR.
PROSITE: PS00237; G PROTEIN RECP F1\_1; 1.
PROSITE: PS0262; G\_PROTEIN\_RECP\_F1\_2; 1.
G-protein coupled receptor; Transmembrane; Glycoprotein.
EXTRACELLULAR (POTENTIAL).
DOMAIN 1 25
TRANSMEM 30 52 1 (POTENTIAL) (POTENTIAL).
DOMAIN 53 65 CYTOPLASMIC (POTENTIAL).
TRANSMEM 66 86 2 (POTENTIAL).
DOMAIN 87 105 EXTRACELLULAR (POTENTIAL).
TRANSMEM 106 127 3 (POTENTIAL).
DOMAIN 128 147 CYTOPLASMIC (POTENTIAL).
TRANSMEM 148 170 4 (POTENTIAL).
DOMAIN 171 191 5 (POTENTIAL).
TRANSMEM 192 217 EXTRACELLULAR (POTENTIAL).
DOMAIN 218 244 CYTOPLASMIC (POTENTIAL).
TRANSMEM 245 268 6 (POTENTIAL).
DOMAIN 269 287 EXTRACELLULAR (POTENTIAL).
TRANSMEM 288 309 7 (POTENTIAL).
DOMAIN 310 341 CYTOPLASMIC (POTENTIAL).
CARBOHYD 4 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 16 16 N-LINKED (GLCNAC. . .) (POTENTIAL).
DISULFID 104 181 BY SIMILARITY.
CONFLICT 188 188 E -> G (IN REF. 2).
SEQUENCE 341 AA; 36927 MW; E85843FE54C1CD94 CRC64;
Query Match 13.2%; Score 66.5; DB 1; Length 341;
STRAIN=Sprague-Dawley; TISSUE=Kidney;

Best Local Similarity 30.2%; Pred. No. 58;  
Matches 32; Conservative 11; Mismatches 38; Indels 25; Gaps 5;

QY 8 GLC-VALSCGAAFLVGSAPFAQPVAALESA-----AEAGAGTANPLGTLNPLKLL 60  
DB 115 GLCPILLGAAMAAERFVGLTFEFSRPAATSRRAWATVGLVWVGAGTIG-----L 163  
QY 61 LESLGI PVNHILEGSKVC-----AELGQAVGAVKALKALLGALT V 102  
DB 164 LPLLGLGRYSVQYPSGNCWFLTGAEGRDVAFG---LMFALLGSVS V 206

RESULT 25  
FOLD\_HUMAN  
ID FOLD\_HUMAN STANDARD; PRT; 455 AA.  
AC Q16676; Q12949;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Forkhead box protein D1 (Forkhead-related protein FKHL8) (Forkhead-related transcription factor 4) (FREAC-4).  
GN FOXD1 OR FKHL8 OR FREAC4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96355467; PubMed=8702877;  
RA Erntsson S., Pierrou S., Hulander M., Cederberg A., Hellqvist M., Carlsson P., Enerbaeck S.;  
RT "Characterization of the human forkhead gene FREAC-4. Evidence for regulation by Wilms' tumor suppressor gene (WT-1) and p53.";  
RL J. Biol. Chem. 271:21094-21099(1996).  
RN [2]  
RP SEQUENCE OF 120-235 FROM N.A.  
RX MEDLINE=95045392; PubMed=7957066;  
RA Pierrou S., Hellqvist M., Samuelsson L., Enerbaeck S., Carlsson P.;  
RT "Cloning and characterization of seven human forkhead proteins: binding site specificity and DNA bending.";  
RL ENDO J. 13:5002-5012(1994)  
CC -1- FUNCTION: Binding of FREAC-3 and FREAC-4 to their cognate sites results in bending of the DNA at an angle of 80-90 degrees.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- SIMILARITY: Contains 1 fork-head domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR ENBL; U59832; AAC50661.1; -  
DR ENBL; U59831; AAC50660.1; -  
DR ENBL; U13222; AAA92039.1; -  
DR PIR; G02738; G02738.  
DR PIR; S51627; S51627.  
DR HSP; Q63245; 2HEP.  
DR TRANSPAC; T02472; -  
DR Gnew; HGNC:3802; FOXD1.  
DR MM; G01091; -  
DR GO; G0003700; Fitranscription factor activity; TAS.  
DR InterPro; IPR001766; TF\_Fork\_head.  
DR Pfam; PF00250; Fork\_head\_1.  
DR PRINTS; PR00053; FORKHEAD.  
DR ProDom; PD000425; TF\_Fork\_head; 1.  
DR SMART; SM00339; FH\_1.  
DR PROSITE; PS00657; FORK\_HEAD\_1; 1.  
DR PROSITE; PS00658; FORK\_HEAD\_2; 1.  
DR PROSITE; PS50039; FORK\_HEAD\_3; 1.  
DR Transcription regulation; DNA-binding; Nuclear protein. KW

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

```
CC CC
CC EMBL; L29767; AAA32995.1; -.
CC DR EMBL; L33907; AAA73948.1; -.
CC DR PIR; S45680; S45680.
CC DR HSSP; F19656; IMZM.
CC DR InterPro; IPR003612; AAI.
CC DR InterPro; IPR00528; Plant LTP.
CC DR Pfam; PF0234; trypt alpha amyl; 1.
CC DR PRINTS; PR00382; LIPIDTRANSF.
CC DR SMART; SMO0499; AAI; 1.
CC DR PROSITE; PS00597; PLANT_LTP; 1.
CC KW Lipid-binding; Transport; signal; Multigene family.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 118 NONSPECIFIC LIPID-TRANSFER PROTEIN D.
FT DISULFID 29 77 POTENTIAL.
FT DISULFID 39 54 POTENTIAL.
FT DISULFID 55 100 POTENTIAL.
FT DISULFID 75 114 POTENTIAL.
SQ SEQUENCE 118 AA; 11937 MW; 53214BCDC4491DFC CRC64;

Query Match 13.1%; Score 66; DB 1; Length 118;
Best Local Similarity 22.7%; Pred. No. 25;
Matches 25; Conservative 18; Mismatches 35; Indels 32; Gaps 3;

Qy 1 MKLAALGLCV-----ALSCSAAAPLVGSAKPVAQPVAALESAAEAGAGTL--- 47
Db 5 MKLAALGLCV-----ALSCSAAAPLVGSAKPVAQPVAALESAAEAGAGTL--- 47
Qy 48 -----ANPLGTLNPLK-----LLSSLGIPVNHLEGSQRC 78
Db 65 ARTPTDQRCRLVGAANALPIINVARAAGLPKACGVNIPYKISTTNC 114

RESULT 27
MP70 MYCKA STANDARD; PRT; 121 AA.
ID Q49614;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Immunogenic protein MP70 analog precursor (Fragment).
OS Mycobacterium kansasii.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1768;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 10268;
RX MEDLINE=97218684; PubMed=9066108;
RA Woolford A.J., Hewinson R.G., Woodward M., Dale J.W.;
RT "Sequence heterogeneity of an mpb70 gene analogue in Mycobacterium kansasii.";
RL FEMS Microbiol. Lett. 148:43-48(1997).
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- SIMILARITY: Contains 1 FAS1 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
FT NON_TER 121 121
SQ SEQUENCE 121 AA; 11930 MW; AD17BD8C67F537E CRC64;

Query Match 13.1%; Score 66; DB 1; Length 121;
Best Local Similarity 32.9%; Pred. No. 25;
Matches 26; Conservative 12; Mismatches 21; Indels 20; Gaps 4;

Qy 4 AALLGLCVALSASSAAFLVG-----SAKPVAQPVAALESAAEAGAGTLA--NPL--- 51
Db 13 AAVGLAVATSTAAADLVGFCADYAAANPSPG--ASVEGMSQVPVAAASNNPMLTT 70
Qy 52 -----GTLNPLKLLSSL 64
Db 71 LITSVSGRLNPQVNLVDYL 89

RESULT 28
CLRC IDEDE STANDARD; PRT; 239 AA.
ID AC P60070;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Chlorate reductase gamma subunit precursor (Chlorate reductase heme
DE subunit).
GN CLRC.
OS Ideonella dechloratans.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Ideonella.
OX NCBI_TaxID=36863;
RN [1];
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RA Danielsson Thorell H., Stenklo K., Karlsson J., Nilsson T.;
RT "A gene cluster for chlorate metabolism in Ideonella dechloratans.";
RL Appl. Environ. Microbiol. 69:5585-5592(2003).
CC -!- FUNCTION: May transfer electrons to the iron-sulfur centers of
CC cfrB.
CC -!- COFACTOR: Binds 1 heme B (iron-protoporphyrin IX) group per
CC molecule (Potential).
CC -!- SUBUNIT: Heterotrimer of alpha, beta and gamma subunits.
CC -!- SUBCELLULAR LOCATION: Periplasmic.
CC -!- BIOTECHNOLOGY: Has potential use in bioremediation of waste sites
CC contaminated with chlorate, such as pulp and paper industry
CC wastewater.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
AC P17225;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Polypyrimidine tract-binding protein 1 (PTB) (Heterogeneous nuclear
DE ribonucleoprotein I) (hnRNP I).
GN PTBP1 OR PTB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 312-341.
RX MEDLINE=92105132; PubMed=1722210;
RA Bothwell A.L.M., Ballard D.W., Philbrick W.M., Lindwall G.,
RA Maher S.E., Bridgett M.M., Jamison S.F., Garcia-Blanco M.A.;
RT "Murine polypyrimidine tract binding protein. Purification, cloning,
RT and mapping of the RNA binding domain."
RL J. Biol. Chem. 266:24657-24663 (1991).
CC -!- FUNCTION: Plays a role in pre-mRNA splicing. Binds to the
CC polypyrimidine tract of introns. May promote the binding of U2
CC snRNP to pre-mRNA.
CC -!- SUBUNIT: Part of a ternary complex containing FUBP2, PTBP1, PTBP2
CC and HNRP1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DOMAIN: The C-terminal 195 amino acids of PTB are sufficient for
CC specific RNA binding.
CC -!- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X52101; CAA36321.1; -.
DR MGD; MGI:97791; Ptbp1.
DR GO; GO:0005634; C:nucleus; IDA.
DR InterPro; IPR006536; HnRNP-L_PTB.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rtm; 4.
DR SMART; SM00360; RRM; 3.
DR TIGRFAMs; TIGR01649; hnRNP-L_PTB; 1.
DR PROSITE; PS00102; RRM; 4.
DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
KW Nuclear protein; RNA-binding; mRNA splicing; Repeat.
FT DOMAIN 58 142 RNA-BINDING (RRM) 1.
FT DOMAIN 183 259 RNA-BINDING (RRM) 2.
FT DOMAIN 335 386 RNA-BINDING (RRM) 3.
FT DOMAIN 450 525 RNA-BINDING (RRM) 4.
FT DOMAIN 315 321 POLY-ALA.
SQ SEQUENCE 527 AA; 56478 MW; F16FDF376010D76A CRC64;
Query Match 13.1%; Score 66; DB 1; Length 527;
Best Local Similarity 26.7%; Pred. No. 94;
Matches 31; Conservative 16; Mismatches 37; Indels 32; Gaps 4;
QY 12 ALSGSSAAAF-----LVGSAKPVAOPVAALSAE-----AGAGTLANPLGTINPLK 58
DB 287 SLDQTAALAFGLSVNVHGAALPLAIPSAAMAAASRTAIPGLAGAGNSVLLNLP 346
QY 59 LLLSLGI-----FVNHLIEGSKQKVAELG--PQAVGAVKAKA 95
DB 347 VTPQSLFLFGVYGDQVKVILFNKKNALVQMDGSOAQLGEPPEAAAREVSA 402
RESULT 30
LEUL_PROMM STANDARD; PRT; 540 AA.
AC Q7TUV5;
DT 15-MAR-2004 (Rel. 43, Created)
```

```
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 2-isopropylmalate synthase (EC 2.3.3.13) (Alpha-isopropylmalate
DE synthase) (Alpha-IPM synthetase).
GN LEUA OR PWT121.
OS Prochlorococcus marinus (strain MIT 9313).
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=74547;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825698; PubMed=12917642;
RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Slegel C., Sullivan S.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinser E.R., Chisholm M.W.;
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT niche differentiation."
RL Nature 424:1042-1047 (2003).
CC -!- FUNCTION: Catalyzes the condensation of the acetyl group of
CC acetyl-CoA with 3-methyl-2-oxobutanoate (2-oxoisovalerate) to form
CC 3-carboxy-3-hydroxy-4-methylpentanoate (2-isopropylmalate).
CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + 3-methyl-2-oxobutanoate + H(2)O =
CC 2-hydroxy-2-isopropylsuccinate + CoA.
CC -!- PATHWAY: Leucine biosynthesis; first step.
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- SIMILARITY: Belongs to the alpha-IPM synthetase / homocitrate
CC synthase family. Leua 1 subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X572098; CAE21296.1; -.
DR HAMAP; MF 01025; -.
DR InterPro; IPR002034; AIPM/Hcit_synth.
DR InterPro; IPR000891; HMGL-like_synth.
DR InterPro; IPR005671; Leua_bact_synth.
DR Pfam; PF00682; HMGL-like; 1.
DR TIGRFAMs; TIGR00973; leua_bact; 1.
DR PROSITE; PS00815; AIPM_HOMOCIT_SYNTH 1; 1.
DR PROSITE; PS00816; AIPM_HOMOCIT_SYNTH 2; 1.
KW Leucine biosynthesis; Transferase; Complete proteome.
SQ SEQUENCE 540 AA; 57970 MW; A68C195F7E8FE7DE CRC64;
Query Match 13.1%; Score 66; DB 1; Length 540;
Best Local Similarity 32.5%; Pred. No. 96;
Matches 26; Conservative 10; Mismatches 34; Indels 10; Gaps 3;
QY 11 VALSCSSAAAFVGSAPVAPVAALSAEAGAGTLANPLGTINPLKLLSLSIGIPVNH 70
DB 412 VQVSCGS-----SLRPTATVILAQEDGQGTAAAVGT--GPVDVCRNALNAGEPNE 462
QY 71 LIEGSKQKVAELGPQAVGAV 90
DB 463 LIETSVKSVTE-GIDAMGEV 481
RESULT 31
NUP1_RAT STANDARD; PRT; 585 AA.
AC P70581; Q9CHE1; Q9QWK7; Q9Z2W7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nucleoporin p58/p45 (Nucleoporin-like protein 1).
GN NUP1.
OS Rattus norvegicus (Rat).
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM P58), SEQUENCE OF 227-237: 317-333;  
 RP 403-422 AND 427-433, ALTERNATIVE SPLICING, FUNCTION, SUBCELLULAR  
 RP LOCATION, IDENTIFICATION IN A COMPLEX WITH NUP62 AND NUP54, AND  
 RP INTERACTION WITH NUPF2.  
 RP TISSUE=Macrophage;  
 RC MEDLINE=96326666; PubMed=8707840;  
 RX Hu T., Guan T., Gerace L.;  
 RA "Molecular and functional characterization of the p62 complex, an  
 RT assembly of nuclear pore complex glycoproteins";  
 RL J. Cell Biol. 134:589-601(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS P45 AND P23), SEQUENCE OF 309-585 FROM  
 RP N.A. (ISOFORM H6), AND TISSUE SPECIFICITY.  
 RC TISSUE=Macrophage;  
 RX MEDLINE=99013879; PubMed=9795236;  
 RA Hu T., Gerace L.;  
 RT "cDNA cloning and analysis of the expression of nucleoporin p45";  
 RL Gene 221:245-253(1998).  
 CC -!- FUNCTION: Component of the nuclear pore complex, a complex  
 CC required for the trafficking across the nuclear membrane.  
 CC -!- SUBUNIT: Component of the p62 complex, a complex composed of  
 CC NUP62, NUP54, and the isoform p58 and isoform p45 of NUP1.  
 CC Isoform p58 interacts with NUPF2. Isoform p58 interacts with SRP1-  
 CC alpha and importin p97 proteins when they are together, but not  
 CC with SRP1-alpha protein alone.  
 CC -!- SUBCELLULAR LOCATION: Central region of the nuclear pore complex,  
 CC within the transporter. Localizes on both cytoplasmic and  
 CC nucleoplasmic sides of the nuclear pore complex near the central  
 CC gated channel.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=4;  
 CC Name=p58;  
 CC IsoId=P70581-1; Sequence=Displayed;  
 CC Name=p45;  
 CC IsoId=P70581-2; Sequence=VSP\_008579, VSP\_008580, VSP\_008581;  
 CC Name=p23;  
 CC IsoId=P70581-3; Sequence=VSP\_008576, VSP\_008577, VSP\_008578,  
 CC VSP\_008582;  
 CC Note=No experimental confirmation available;  
 CC Name=H6;  
 CC IsoId=P70581-4; Sequence=VSP\_008583, VSP\_008584;  
 CC Note=No experimental confirmation available;  
 CC -!- TISSUE SPECIFICITY: Expressed in liver.  
 CC -!- DOMAIN: Contains P-G repeats.  
 CC -!- PTM: O-glycosylated.  
 CC -!- SIMILARITY: Belongs to the NUP1 family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U63839; AAC2789.1; -;  
 CC DR EMBL; AF000898; AAC82539.1; -;  
 CC DR EMBL; AF000900; AAC82318.1; -;  
 CC DR EMBL; AF000901; AAC82319.1; -;  
 KW Transport; Nuclear protein; Repeat; Coiled coil; Glycoprotein;  
 KW Alternative splicing.  
 FT DOMAIN 7 565 14 X 2 AA REPEATS OF F-G.  
 FT REPEAT 7 8 1.  
 FT REPEAT 30 31 2.  
 FT REPEAT 42 43 3.  
 FT REPEAT 61 62 4.  
 FT REPEAT 66 67 5.  
 FT DOMAIN 242 262 COILED COIL (POTENTIAL).  
 FT DOMAIN 300 367 COILED COIL (POTENTIAL).  
 FT

FT REPEAT 474 475  
 FT REPEAT 478 479  
 FT REPEAT 499 500  
 FT REPEAT 505 506  
 FT REPEAT 515 516  
 FT REPEAT 517 518  
 FT REPEAT 531 532  
 FT REPEAT 554 555  
 FT REPEAT 564 565  
 FT VARSPLIC 328 337  
 FT  
 FT VARSPLIC 466 468  
 FT VARSPLIC 469 585  
 FT VARSPLIC 224 245  
 FT VARSPLIC 531 535  
 FT VARSPLIC 536 585  
 FT VARSPLIC 1 327  
 FT VARSPLIC 466 471  
 FT VARSPLIC 472 585  
 FT SEQUENCE 585 AA; 59265 MW; 4DSF5C2744A75C3C CRC64;  
 Query Match 13.1%; Score 66; DB 1; Length 585;  
 Best Local Similarity 27.5%; Pred. No. Id=02;  
 Matches 30; Conservative 11; Mismatches 48; Indels 20; Gaps 4;  
 Qy 12 ALSGSSAAAFVLS-----SAKPVAPQV-----AALESAAEAGAGTLANPLCTL 54  
 Db 85 ATTSASTTGSLGFPKPAASATPPALPTVTTASGLTSSALTSAPASTGFTLNNLGA- 143  
 Qy 55 NPKLKLLSLGIPVNHLEGSOKVAELGPQAVGAV--KALKALGALT 101  
 Db 144 TPATTTAASGTLGGLAGLAGLGGSLFQSGNTATSGLGQNALSLSLGSTAT 192  
 RESULT 32  
 EUTC\_RHOER STANDARD; PRT; 257 AA.  
 ID EUTC\_RHOER  
 AC Q59782;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Ethanolamine ammonia-lyase light chain (EC 4.3.1.7) (Ethanolamine  
 DE ammonia-lyase small subunit).  
 GN EUTC.  
 OS Rhodococcus erythropolis.  
 OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Nocardiaceae; Rhodococcus.  
 OX NCBI\_TaxID=1833;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=N786721;  
 RX MEDLINE=94349216; PubMed=8069783;  
 RA De Mot R., Nagy I., Schoofs G., Vanderleyden J.;  
 RT "Sequence of Rhodococcus gene cluster encoding the subunits of  
 RT ethanolamine ammonia-lyase and an APC-like permease";  
 RL Can. J. Microbiol. 40:403-407(1994).  
 CC -!- CATALYTIC ACTIVITY: Ethanolamine = acetaldehyde + NH(3).  
 CC -!- COFACTOR: Adenosylcobalamin.  
 CC -!- PATHWAY: Ethanolamine utilization.  
 CC -!- SUBUNIT: Heterodimer of a small and a large subunit (By  
 CC similarity).  
 CC -!- SIMILARITY: Belongs to the eutC family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L24492; AAC37137.1; -;  
DR HAMAP; MF 00601; -; 1.  
DR Lyase; Cdbalt.  
SQ SEQUENCE 257 AA; 26963 MW; E6E3FA138P49C91B CRC64;  
  
Query Match 13.0%; Score 65.5; DB 1; Length 257;  
Best Local Similarity 25.3%; Pred. No. 55;  
Matches 25; Conservative 18; Mismatches 41; Indels 15; Gaps 4;  
  
QY 17 SAAFLVGSAPVAPVAALLES-----AAAGAGTIANPLGTNPFLKLLSSIGIPV 68  
DB 130 SALVTALGERSYIAPVATNARVALGDHIAAGVQTAIVLIGE-RPGLSVADSVGIYL 188  
QY 69 NHL-----IEGSKCVKVEL-GPQAVGAVKALKALIGALT 101  
DB 189 THLPVRGRTDARNCSINSHPPPEGLGYEQARVVVLGVT 227  
  
RESULT 33  
COBT\_RHILO STANDARD; PRT; 336 AA.  
AC Q98KN9;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Nicotinate-nucleotide-dimethylbenzimidazole phosphoribosyltransferase  
DE (EC 2.4.2.21) (NN:DBI PRT) (N(1)-alpha-phosphoribosyltransferase).  
GN COBT OR MUR1389.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MAFF303099;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RA "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
RT Mesorhizobium loti";  
RL DNA Res. 7:331-338(2000).  
CC -!- FUNCTION: Catalyzes the synthesis of alpha-ribazole-5'-phosphate  
CC from nicotinate mononucleotide (NAMN) and 5,6-  
CC dimethylbenzimidazole (DMB).  
CC -!- CATALYTIC ACTIVITY: Beta-nicotinate D-ribonucleotide +  
CC ribosyl-5,6-dimethylbenzimidazole = nicotinate + N(1)-(5-phospho-alpha-D-  
CC ribosyl)-5,6-dimethylbenzimidazole.  
CC -!- PATHWAY: Cobalamin biosynthesis.  
CC -!- SIMILARITY: Belongs to the cobT family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AP002997; BAB48775.1; -;  
DR HSSP; Q05603; 1DOS.  
DR HAMAP; MF 00230; -; 1.  
DR InterPro; IPR003200; NN:DBI\_PRT.  
DR InterPro; IPR008281; NN:DBI\_PRT\_sub.

DR Pfam; PF02277; DBI\_PRT; 1.  
DR ProDom; PD009438; NN:DBI\_PRT; 1.  
KW Cobalamin biosynthesis; Porphyrin biosynthesis; Transferase;  
KW Glycosyltransferase; Complete proteome.  
FT ACT SITE 304 BASE (BY SIMILARITY).  
SQ SEQUENCE 336 AA; 34408 MW; F9824D6693745932 CRC64;  
  
Query Match 13.0%; Score 65.5; DB 1; Length 336;  
Best Local Similarity 26.9%; Pred. No. 70;  
Matches 35; Conservative 12; Mismatches 44; Indels 39; Gaps 5;  
  
QY 7 LGICVALSCSSAAAFVLSAK-----PVAQPVAALLESAAEAGAGTIANPLGTILN 55  
DB 164 IGNTTAAAI-SAAALFGGGAEKMTGRGTGVDAGLKKVKVAIEAGLKRHAALADPLGVA 223  
QY 56 PL-----KLLLSLIGIPVNLHIGSKCVKVELGPQA-----VGAV 90  
DB 224 ALGRELALFGATLAARHLGIPV--LLDGF-VCTRAAAPLARLHPTGLSHITIAHVSNE 280  
QY 91 KALKALIGAL 100  
DB 281 SGHRLLEAL 290  
  
RESULT 34  
YO84\_MYCTU STANDARD; PRT; 491 AA.  
AC O53209;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical protein RV2484C/M12557.  
GN RV2484C OR M12557 OR M12557.40C.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37RV;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaiia F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;  
RA "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence";  
RL Nature 393:537-544(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CDC 1551 / Oshkosh;  
RX MEDLINE=22206494; PubMed=12218036;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Umayam L.A., Emolaeva W., Salzberg S.L.,  
RA Decher A., Utterback T., Weidman J., Venter J.C., Fraser C.M.;  
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;  
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
RT laboratory strains";  
RL J. Bacteriol. 184:5479-5490(2002).  
CC -!- SIMILARITY: Belongs to the UPF0089 family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



CC -----  
DR EMBL: AL021246; CAAL6061.1; --  
DR EMBL: AE007092; AAK46861.1; --  
DR PIR: C70868; C70868.  
DR TIGR: W12557; --  
DR Tuberculist; RV2484c; --  
DR InterPro; IPR004255; UPF0089.  
DR Pfam; PF03007; UPF0089; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 491 AA; 52309 MW; C2ACCE458ADP82DD CRC64;  
  
Query Match 13.0%; Score 65.5; DB 1; Length 491;  
Best Local Similarity 36.7%; Pred. No. 98;  
Matches 22; Conservative 8; Mismatches 23; Indels 7; Gaps 2;  
  
QY 4 AALLGICVALSCSSAAAFVLSAKFVAQVVALESAAEAGAGTIANPLGTINPL 57  
DB 281 AYLGLGCGALRRYHEALGVISTLPMVAVV-NLRAEGDAAGNGFTGVNLAAPVGTIDPV 339  
  
RESULT 35  
TRD2 FALSO  
ID TRD2 FALSO STANDARD; PRT; 344 AA.  
AC Q8XS00;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Anthranilate phosphoribosyltransferase 2 (EC 2.4.2.18).  
GN TRPD2 OR RSP0681 OR RSP01768.  
OS Ralstonia solanacearum (Pseudomonas solanacearum).  
OG Plasmid megaplasmid.  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Burkholderiaceae; Ralstonia.  
OX NCBI\_TaxID=305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21691879; PubMed=11823852;  
RA Salanoubat M., Gsinn S., Artiguenave F., Gouzy J., Mangenot S.,  
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,  
RA Chandler M., Choise N., Claude-Renaud C., Cunrac S., Demange N.,  
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,  
RA Siguer P., Thebaud P., Whalen M., Wincker F., Levy M.,  
RA Weissenbach J., Boucher C.A.;  
RA "Genome sequence of the plant pathogen Ralstonia solanacearum";  
RL Nature 415:497-502(2002).  
CC -!- CATALYTIC ACTIVITY: Anthranilate + phosphoribosyl-diphosphate =  
CC N-5-phosphoribosyl-anthranilate + diphosphate.  
CC -!- PATHWAY: Tryptophan biosynthesis; second step.  
CC -!- SIMILARITY: Belongs to the anthranilate phosphoribosyltransferase  
CC family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AL646080; CAD17832.1; --  
DR HAMAP; MF\_02211; 1.  
DR InterPro; IPR005940; Ant\_phospho\_trans.  
DR InterPro; IPR000312; Glyco\_transf\_3.  
DR Pfam; PF02885; Glycos\_transf\_3N; 1.  
DR Pfam; PF00591; Glycos\_transf\_3; 1.  
DR ProDom; PD001864; Glyco\_transf\_3; 1.  
DR TIGRfam; TIGR01245; trpd; 1.  
KW Tryptophan biosynthesis; Transferase; Glycosyltransferase; Plasmid;  
KW Complete proteome.  
SQ SEQUENCE 344 AA; 35977 MW; 65074865254E2F2 CRC64;  
  
Query Match 12.9%; Score 65; DB 1; Length 344;

Best Local Similarity 27.6%; Pred. No. 79;  
Matches 27; Conservative 18; Mismatches 29; Indels 24; Gaps 5;  
  
QY 16 SSAAAFVLSAKFVAQVVALESAAEAGAGTIANPLGTINPLKLLSLGIPVNHLS 74  
DB 93 STCAAFIVAGAGVEV-----AKHGNRAUSSKSGAAD-----VLSALGVNLDQTPAD 138  
  
QY 75 SOKCVAEIG-----POAVGAVKALKAL---LGLATVTF 103  
DB 139 IERCIAEAGIGFMPAPTHHPALKQLMPFVRVDLATRTIF 176  
  
RESULT 36  
YM96 SYNEL  
ID YM96 SYNEL STANDARD; PRT; 399 AA.  
AC Q8DGM0;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical UPF0272 protein t112296.  
GN T112296.  
OS Synechococcus elongatus (Thermosynechococcus elongatus).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.  
OX NCBI\_TaxID=32046;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=BP-1;  
RX MEDLINE=2225144; PubMed=12240834;  
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,  
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,  
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,  
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;  
RA "Complete genome structure of the thermophilic cyanobacterium  
RT Thermosynechococcus elongatus BP-1";  
RL DNA Res. 9:123-130(2002).  
CC -!- SIMILARITY: Belongs to the UPF0272 family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AP005376; BAC09848.1; --  
DR HAMAP; MF\_01074; 1.  
DR InterPro; IPR002822; DUF111.  
DR Pfam; PF01969; DUF111; 1.  
DR ProDom; PD018127; DUF111; 1.  
DR Hypothetical protein; Complete proteome.  
KW SEQUENCE 399 AA; 43522 MW; 9742A63BD0C9CE12 CRC64;  
  
Query Match 12.9%; Score 65; DB 1; Length 399;  
Best Local Similarity 34.0%; Pred. No. 90;  
Matches 34; Conservative 10; Mismatches 34; Indels 22; Gaps 7;  
  
QY 12 ALSCSSAAAFVLSAKFVAQVVALESAAEAGAGTIANPLGTINPLKLLSLGIPVNH 71  
DB 200 ALVCALSAQF-----GAPPANTLQRVG-LGAGTQELPLNL--LRLLWLTGVPAPPE-- 247  
  
QY 72 IEGSQKCVAEI-----GFOAVG-AVKALKALGLATVTF 103  
DB 248 -TGTAKTIVELQTLDDMPFQALSYACEOLYA-AGALEVF 285  
  
RESULT 37  
A2AC HUMAN  
ID A2AC HUMAN STANDARD; PRT; 462 AA.  
AC P18825; P35369; Q9HB49;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)



RL J. Bacteriol. 173:6837-6843 (1991).  
RP [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12;  
RC MEDLINE=92184121; PubMed=1544582;  
RA Kalman M., Murphy H., Cashel M.;  
RT "The nucleotide sequence of recG, the distal spo operon gene in  
RL Escherichia coli K-12.";  
RL Gene 110:95-99(1992).  
RN [3]  
RN SEQUENCE FROM N.A.  
RP STRAIN-K12 / MG1655;  
RC MEDLINE=93315143; PubMed=7686882;  
RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;  
RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli  
RT genome: organizational symmetry around the origin of replication.";  
RL Genomics 16:551-561(1993).  
RN [4]  
RN FUNCTION.  
RP MEDLINE=95045415; PubMed=7957087;  
RA Whitby M.C., Vincent S.D., Lloyd R.G.;  
RT "Branch migration of Holliday junctions: identification of RecG  
RT protein as a junction specific DNA helicase.";  
RL EMBO J. 13:5220-5228(1994).  
RN [5]  
RN FUNCTION.  
RP MEDLINE=95292979; PubMed=7774596;  
RA Hong X., Cadwell G.W., Kogoma T.;  
RT "Escherichia coli RecG and RecA proteins in R-loop formation.";  
RL EMBO J. 14:2385-2392(1995).  
CC -!- FUNCTION: Critical role in recombination and DNA repair. Help  
CC process Holliday junction intermediates to mature products by  
CC catalyzing branch migration. Has a DNA unwinding activity  
CC characteristic of a DNA helicase with a 3' to 5' polarity. recG  
CC unwind branched duplex DNA (Y-DNA). Has a role in constitutive  
CC stable DNA replication (cedR) and R-loop formation.  
CC -!- SIMILARITY: Belongs to the helicase family. RecG subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X59550; CAA42123.1; -.  
DR EMBL; M64367; AAR24513.1; -.  
DR EMBL; L10328; AAR62005.1; -.  
DR EMBL; AE000042; AAC76676.1; -.  
DR PIR; JH0265; JH0265.  
DR EcoGene; EG10829; recG.  
DR InterPro; IPR001410; DEAD.  
DR InterPro; IPR001650; Helicase\_C.  
DR InterPro; IPR008994; Nucleic\_acid\_OB.  
DR InterPro; IPR004509; RecG.  
DR InterPro; IPR004365; trna\_anti.  
DR Pfam; PF00270; DEAD; 1.  
DR Pfam; PF00271; helicase\_C; 1.  
DR Pfam; PF01336; trna\_anti; 1.  
DR SMART; SM00487; DEXDC; 1.  
DR SMART; SM00490; HELIC; 1.  
DR TIGFams; TIGR00643; recG; 1.  
DR HydroLase; Helicase; DNA repair; ATP-binding; DNA recombination;  
KW DNA-binding; Complete proteome.  
FT NP\_BIND 296 303 ATP (POTENTIAL).  
FT SITE 397 400 DEQH BOX.  
FT SEQUENCE 693 AA; 76430 MW; 7826143A8F4292A2 CRC64;  
Query Watch 12.9%; Score 65; DB 1; Length 693;  
Best Local Similarity 30.2%; Pred.No. 1.5e-02;  
Matches 26; Conservative 10; Mismatches 49; Indels 6; Gaps 2;

QY 23 VGSAPVPAQPVAALESAAEAGAGTLANPLGTL-----NPLKLLSSLGIPVNHLEGSQK 77  
Db 298 VSGGKTVIAALAAALRAIAHGKQVALMAPELLAQHANFENWFAPIGIEVGWL-AGKQK 356  
QY 78 CVAELGFPQAVGAVKALKALLGALTTF 103  
Db 357 GKARLAQQAIAASGQVQMTIVGTHAIF 382  
RESULT 39  
ANK3 HUMAN  
ID ANK3\_HUMAN STANDARD; PRT; 4377 AA.  
AC Q12955;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Ankyrin 3 (ANK-3) (Ankyrin G).  
GN ANK3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain stem;  
RC MEDLINE=95138209; PubMed=7836469;  
RA Kordeli E., Lambert S., Bennett V.;  
RT "Ankyrin, A new ankyrin gene with neural-specific isoforms localized  
RT at the axonal initial segment and node of Ranvier.";  
RL J. Biol. Chem. 270:2352-2359(1995).  
CC -!- FUNCTION: Membrane-cytoskeleton linker. The neural-specific  
CC isoforms may participate in the maintenance/targeting of ion  
CC channels and cell adhesion molecules at the nodes of Ranvier and  
CC axonal initial segments.  
CC -!- SUBUNIT: Neural-specific isoforms may be a constituent of a  
CC neurofascin/NRCAM/ankyrin G complex.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named isoforms=1;  
CC Comment=A number of isoforms are produced;  
CC Name=1; Synonyms=480-kDa isoform;  
CC IsoId=Q12955-1; Sequence=Displayed;  
CC -!- TISSUE SPECIFICITY: Expressed in brain and other tissues. Isoform  
CC 1 is neural-specific.  
CC -!- SIMILARITY: Contains 23 ANK repeats.  
CC -!- SIMILARITY: Contains 1 death domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U13616; AAA64834.1; -.  
DR PIR; A55575; A55575.  
DR HSSP; P55273; 1B18.  
DR Genew; HGNC:494; ANK3.  
DR MIM; 600465;  
DR GO; GO:0006605; P; protein targeting; NAS.  
DR InterPro; IPR002110; ANK.  
DR InterPro; IPR000488; Death.  
DR InterPro; IPR000906; ZU5.  
DR Pfam; PF00023; ank; 24.  
DR Pfam; PF00531; death; 1.  
DR Pfam; PF00791; ZU5; 1.  
DR PRINTS; PR01415; ANKYRIN.  
DR SMART; SM00248; ANK; 21.  
DR SMART; SM00005; DEATH; 1.  
DR SMART; SM00218; ZU5; 1.  
DR PROSITE; PS50088; ANK\_REPEAT; 21.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
DR PROSITE; PS50017; DEATH\_DOMAIN; 1.

```

KW Cytoskeleton; Alternative splicing; Repeat; ANK repeat.
FT REPEAT 73 102 ANK 1.
FT REPEAT 106 135 ANK 2.
FT REPEAT 139 168 ANK 3.
FT REPEAT 172 201 ANK 4.
FT REPEAT 203 230 ANK 5.
FT REPEAT 234 263 ANK 6.
FT REPEAT 267 296 ANK 7.
FT REPEAT 300 329 ANK 8.
FT REPEAT 333 362 ANK 9.
FT REPEAT 365 395 ANK 10.
FT REPEAT 399 428 ANK 11.
FT REPEAT 432 461 ANK 12.
FT REPEAT 465 494 ANK 13.
FT REPEAT 498 527 ANK 14.
FT REPEAT 531 560 ANK 15.
FT REPEAT 564 593 ANK 16.
FT REPEAT 597 626 ANK 17.
FT REPEAT 630 659 ANK 18.
FT REPEAT 663 692 ANK 19.
FT REPEAT 696 725 ANK 20.
FT REPEAT 729 758 ANK 21.
FT REPEAT 762 791 ANK 22.
FT REPEAT 795 825 ANK 23.
FT REPEAT 828 857 ANK 24.
FT DOMAIN 1519 1898 SER-RICH.
FT DOMAIN 4030 4174 DEATH.
SQ SEQUENCE 4377 AA; 480399 MW; P42379B55768B684 CRC64;

Query Match 12.9%; Score 65; DB 1; Length 4377;
Best Local Similarity 26.2%; Pred. No. 7.7e+02;
Matches 27; Conservative 17; Mismatches 35; Indels 24; Gaps 4;

QY 16 SSAAPFLVGSAPVAPVAA-----LESAEAGAGATLANPL-----GTLNPLK 58
DB 1665 TSAAPFLISPLKSVSPVYSKRVVIVSSAKITWASLSLSPVKQMPGHAVALVNGSI SPLK 1724
QY 59 LLLSLGIPVNHLEIGSKVQAEIQLPQAVGAVKALKALLGALT 101
DB 1725 YASSST-----LINGC-KATATLQEKISSATNSVSVVSAAT 1760

RESULT 40
CYCM BRAJA
ID CYCM BRAJA STANDARD; PRT; 184 AA.
AC P30323;
DT 01-APR-1993 (Rel. 25, Created)
DI 01-APR-1993 (Rel. 25, Last sequence update)
DI 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytochrome c homolog.
GN CYCM OR BLR1423.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110SPC4;
RC MEDLINE=92041558; PubMed=1657867;
RA Bott M., Ritz D., Hennecke H.;
RA "The Bradyrhizobium japonicum cym gene encodes a membrane-anchored
RT homolog of mitochondrial cytochrome c.";
RL J. Bacteriol. 173:6766-6772(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RC MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamiasawa K., Uchiumi T.;
RA Sasamoto S., Watanabe A., Ideawa K., Iiguchi M., Kawashima K.;
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RL Bradyrhizobium japonicum USDA110.";
DNA Res. 9:189-197(2002).

```

```

CC -!- FUNCTION: MAY BE INVOLVED IN ELECTRON TRANSFER FROM B-C1 COMPLEX
CC TO AA3.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M77189; AAA26198.1; -.
CC DR EMBL; AF005940; BAC46688.1; -.
CC DR PIR; A41331; A41331.
CC DR HSSP; P00004; 1WEJ.
CC DR InterPro; IPR003088; Cyt_C1.
CC DR InterPro; IPR002327; Cyt_C1AB.
CC DR InterPro; IPR000345; Cyt_heme_BS.
CC DR Pfam; PF00034; cytochrome c; 1.
CC DR PRINTS; PR00604; CYTCRMECIAB.
CC DR PRODOM; PD000375; Cyt_C1AB; 1.
CC DR PROSITE; PS00190; CYTOCHROME C; 1.
CC DR Electron transport; Heme; Transmembrane; Signal-anchor;
KW Complete proteome.
FT DOMAIN 1 10 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 11 31 SIGNAL-ANCHOR (POTENTIAL).
FT DOMAIN 32 184 PERIPLASMIC (POTENTIAL).
FT BINDING 84 84 HEME (COVALENT) (BY SIMILARITY).
FT BINDING 87 87 HEME (COVALENT) (BY SIMILARITY).
FT METAL 88 88 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT METAL 151 151 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 184 AA; 19098 MW; 2ECDFA564389824 CRC64;

Query Match 12.8%; Score 64.5; DB 1; Length 184;
Best Local Similarity 26.9%; Pred. No. 50;
Matches 28; Conservative 14; Mismatches 39; Indels 23; Gaps 5;

QY 3 LAALGLCVALSASSAAFLVGSAPVAPQ-----VALESAA---EAGAGTLANPLGLAP 56
DB 10 LGAVLGTCILLLVTSFTANALFSPKPEKGFPIAVKEDAGHGKGGGAAAAA-S-----EP 64
QY 57 LKLLSLGIPVNHLEIGSKQCV-----ELGPQAVGAV 90
DB 65 IEKLLQTASVEKG--AAAAKCGAGCTFEKGGPNRVGNLYGVV 106

Search completed: April 5, 2004, 14:34:35
Job time : 19 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 5, 2004, 14:28:38 ; Search time 39 Seconds  
(without alignments)  
841.382 Million cell updates/sec

Title: US-09-997-428-408

Perfect score: 502

Sequence: 1 MKLAALLGLCVALLSCSSAAA.....QAVGAVKALKALIGALTVEFG 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database :

SPTREMBL 25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mbc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_virus:\*

16: sp\_bacteriophage:\*

17: sp\_archaeophages:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	225	44.8	94	11	Q8CJC6 mus musculus
2	88	17.5	255	16	Q82LN6
3	83.5	16.6	281	16	Q8F5T3
4	82.5	16.4	247	16	Q9K3G4 streptomyces
5	82.5	16.4	694	16	Q98CR1 rhizobium 1
6	81	16.1	496	10	Q9SU83
7	80	15.9	148	16	Q8XW23 ralstonia s
8	79	15.7	136	16	Q82B50 streptomyces
9	77.5	15.4	229	16	Q7WJ66 bordetella
10	77.5	15.4	229	16	Q7WA34
11	77.5	15.4	229	16	Q7VYAL bordetella
12	77.5	15.4	601	16	Q86691 streptomyces
13	77	15.3	453	17	Q8TH10 methanobacterium
14	76	15.1	355	16	Q8UA14
15	76	15.1	699	16	Q8BL0 caulobacter
16	76	15.1	889	16	Q8GX4 bifidobacter

17	75.5	15.0	459	16	Q9F2L6 streptomyces
18	75.5	15.0	496	16	Q8G3S5
19	75.5	15.0	544	17	Q9YEG2 aeropyrum p
20	75.5	15.0	667	16	Q8XRM8 ralstonia s
21	75	14.9	299	16	Q89XU3
22	75	14.9	355	2	Q8RPK5
23	75	14.9	370	16	Q89H46
24	75	14.9	463	11	Q8G313
25	74	14.7	611	16	Q9L0K1
26	73.5	14.6	119	5	Q86D11
27	73.5	14.6	331	17	Q9HN19
28	73.5	14.6	358	16	Q8NQ28
29	73.5	14.6	389	16	Q8EJ01 shewanella
30	73.5	14.6	1468	5	Q9GUB5
31	73	14.5	317	16	Q98D48
32	73	14.5	429	16	Q825M1
33	73	14.5	502	16	Q915I2
34	73	14.5	542	10	Q8GTA1
35	73	14.5	584	16	Q9AAR5
36	73	14.5	620	16	Q7WL37
37	73	14.5	620	16	Q7W7Q0
38	73	14.5	620	16	Q7VV87
39	73	14.5	1279	5	Q46099
40	72.5	14.4	240	2	Q84CV5
41	72.5	14.4	373	5	Q7YWC1
42	72.5	14.4	384	16	Q88KB1
43	72.5	14.4	421	2	Q93I77
44	72.5	14.4	482	5	Q9BHF9
45	72.5	14.4	616	16	Q7W2M5
46	72.5	14.4	3165	2	Q93921
47	72	14.3	236	5	Q26351
48	72	14.3	413	16	Q98NR3
49	72	14.3	528	16	P71969 mycobacteri
50	72	14.3	528	16	Q7TY52
51	72	14.3	531	16	Q9RJ14
52	71.5	14.2	343	16	Q88WU3
53	71.5	14.2	462	16	Q9A3U4
54	71.5	14.2	474	16	Q89FZ0
55	71.5	14.2	520	10	Q94E69
56	71.5	14.2	694	16	Q8KBL2
57	71.5	14.2	887	16	Q8UBS5
58	71	14.1	275	16	Q82U00
59	71	14.1	315	16	Q89P63
60	71	14.1	322	10	Q851R4
61	71	14.1	337	16	Q89TF3
62	71	14.1	341	2	Q9AMX3
63	71	14.1	409	10	Q8S5W5
64	71	14.1	422	16	Q82UN5
65	71	14.1	440	16	Q911Q3
66	71	14.1	448	16	Q826U4
67	71	14.1	947	16	Q8RJP8
68	70.5	14.0	260	2	Q93EW4
69	70.5	14.0	326	12	Q9DPB4
70	70.5	14.0	353	16	Q8QX5
71	70.5	14.0	412	11	Q99J73
72	70.5	14.0	668	10	Q23323
73	70.5	14.0	691	16	Q8FTB0
74	70	13.9	125	2	Q33526
75	70	13.9	334	16	Q98B15
76	70	13.9	366	16	Q88I11
77	70	13.9	403	16	Q8FWQ3
78	70	13.9	431	10	Q23028
79	70	13.9	452	10	Q9SCQ6
80	70	13.9	466	16	Q82NJO
81	70	13.9	477	16	Q9XEN9
82	70	13.9	491	17	Q9HFN5
83	70	13.9	602	16	Q8SVK1
84	70	13.9	689	16	Q98EK7
85	69.5	13.8	88	5	Q86D13
86	69.5	13.8	102	5	Q86D12
87	69.5	13.8	244	16	P74128 synechocyst
88	69.5	13.8	305	11	Q8BM41
89	69.5	13.8	362	16	Q88QE2

90 69.5 13.8 387 16 Q88L86  
91 69.5 13.8 396 16 Q8RVR2  
92 69.5 13.8 409 16 Q8Y2W6  
93 69.5 13.8 423 16 Q7UL14  
94 69.5 13.8 426 16 Q05907  
95 69.5 13.8 427 2 Q9FB37  
96 69.5 13.8 455 16 Q8ZNL3  
97 69.5 13.8 455 16 Q8Z599  
98 69.5 13.8 495 16 Q8NTN6  
99 69.5 13.8 567 16 Q98M48  
100 69.5 13.8 636 16 Q98SQ7

ALIGNMENTS

RESULT 1  
Q8CUC6 PRELIMINARY; PRT; 94 AA.  
AC Q8CUC6  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE UGRP2 type B.  
GN SCGB3A1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22326074; PubMed=12438750;  
RA Niimi T., Copeland N.G., Gilbert D.J., Jenkins N.A., Srisodhai A.,  
Zimonjic D.B., Keck-Waggoner C.L., Popescu N.C., Kimura S.;  
RT "Cloning, expression, and chromosomal localization of the mouse gene  
(Scgb3a1, alias Ugrp2) that encodes a member of the novel uteroglobin-  
related protein gene family."  
RL Cytogenet. Genome Res. 97:120-127(2002).  
DR EMBL; AF313457; AAN62328.1; --  
DR MGB; MGI:1915912; SCGB3A1.  
SQ SEQUENCE 94 AA; 9578 MW; 7C84B908A6365B59 CRC64;

Query Match 44.8%; Score 225; DB 11; Length 94;  
Best Local Similarity 57.3%; Pred. No. 2e-13;  
Matches 51; Conservative 14; Mismatches 20; Indels 4; Gaps 2;  
QY 17 SAARFLVGSAPVAPVAALAEAGAGTLAN-PLGTLPKLLSSLGIPVNHLEGS 75  
DB 9 SVAFMDSLAKPAVEPVAAALAPAAEAVAGVPSLPDLHLAIFILASGIFLDELIEGS 68  
QY 76 QKCVAEGLPQAVGAVKALKALGATVFG 104  
DB 69 RKCVTGLGPEAVGAV---KSLGLVLTWFG 94

RESULT 2  
Q82LN6 PRELIMINARY; PRT; 255 AA.  
AC Q82LN6  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative methionine aminopeptidase.  
GN SAV1974.  
OS Streptomyces avermitilis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomyces; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=33903;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=WA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
RX MEDLINE=21477403; PubMed=11572948;  
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,

Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,  
Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
RT "Genome sequence of an industrial microorganism Streptomyces  
avermitilis; deducing the ability of producing secondary  
metabolites."  
Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=WA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
RX MEDLINE=2269306; PubMed=12692562;  
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
Sakaki Y., Hattori M., Omura S.;  
RT "Complete genome sequence and comparative analysis of the industrial  
microorganism Streptomyces avermitilis."  
Nat. Biotechnol. 21:526-531(2003).  
DR EMBL; AP005029; BAC69685.1; --  
DR GO; GO:0004239; F:methionyl aminopeptidase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR000994; Peptidase M24.  
DR InterPro; IPR002467; Pept M24A\_MAP1.  
DR InterPro; IPR001714; Pept\_M24\_MAP.  
DR Pfam; PF00557; Peptidase\_M24\_1.  
DR PRINTS; PR00599; MAPEPTIDASE.  
DR TIGRFAMs; TIGR00500; met\_pcase\_1; 1.  
KW Aminopeptidase; Complete proteome.  
SQ SEQUENCE 255 AA; 26727 MW; 0C41CD3010F855A3 CRC64;  
Query Match 17.5%; Score 88; DB 16; Length 255;  
Best Local Similarity 31.6%; Pred. No. 2.7; Indels 32; Gaps 6;  
Matches 31; Conservative 16; Mismatches 19;  
QY 7 LGICVA---LSCSSAAFLVGSAPVAPVALE-----SAAEAGA--GTLANPL 51  
DB 99 LDLAVARGVGAADAISFLVGKARP-ABSVAIMEATERALAAGIAAKPGARIGDLSHAI 157  
QY 52 GTLNPLKLLSSLGIPVNHLEGSQKVAELGPQAVGA 89  
DB 158 GT-----VLKAGYEPIN-----TEFGHGIGS 179  
RESULT 3  
Q8FST3 PRELIMINARY; PRT; 281 AA.  
AC Q8FST3  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Conserved hypothetical protein.  
GN CE0299.  
OS Corynebacterium efficiens.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
OX NCBI\_TaxID=152794;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;  
RA Kawarayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,  
Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,  
Usuda Y., Sugimoto S.;  
RT "The entire genomic sequence of Corynebacterium efficiens YS-314."  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP005215; BAC17109.1; --  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 281 AA; 28188 MW; 3143B69C35048C2D CRC64;  
Query Match 16.8%; Score 83.5; DB 16; Length 281;  
Best Local Similarity 30.1%; Pred. No. 7.7; Indels 39; Gaps 7;  
Matches 40; Conservative 9; Mismatches 45;  
QY 3 LAALLGLCV-----ALSCSSAAFLVGSAPVAPVALEGAEE-----AGAGTL 47  
DB 106 LAGFLGLICIGNLAGAPMADMDHALHTTGTSGAGPTTVALTAARRVRSGGGAVAL 165

```

QY 48 ANP-----LGTL-----NPKLLSLGIPVNHLEGS-----QKVAEL-QPQAV 87
:      :      :      :      :      :      :      :      :      :
Db 166 IDAPMDLQRLGTIWEVBEGHPIVRLLDQI-----KHLREAQHRHQAAQLOGPQAT 221
:      :      :      :      :      :      :      :      :      :
QY 88 GAVKALKALLGAL 100
:      :      :      :      :      :      :      :      :      :
Db 222 AVILLALLPLAGVL 234
:      :      :      :      :      :      :      :      :      :

RESULT 4
Q9K3G4 PRELIMINARY; PRT; 247 AA.
ID Q9K3G4
AC Q9K3G4
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein SC01277.
GN SC01277 OR 2SG18.24.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteris; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL, AL939108; CAB99157.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 247 AA; 22950 MW; E832ABD76146881B CRC64;

Query Match 16.4%; Score 82.5; DB 16; Length 247;
Best Local Similarity 31.8%; Pred.No. 8.3;
Matches 34; Conservative 10; Mismatches 40; Indels 23; Gaps 5;

QY 17 SAAAPFLVGSAPK-----VAQPVAALESAAAG-----ACTLANPLGTI-----NPKL 59
:      :      :      :      :      :      :      :      :      :
Db 61 NALAAGAVNPATLAGTTPPLAALGGIGAGNPLAGLAGAGACPLNTAGAAQNPLAA 120
:      :      :      :      :      :      :      :      :      :

QY 60 LLSSL--GIPVNHLEGSQKVAELGQVAVGAVKALKALGALTIVFG 104
:      :      :      :      :      :      :      :      :      :
Db 121 LTGAAGGNNPLAALGGAGNPLAALG-----GAANPLAAVGGAGALG 163
:      :      :      :      :      :      :      :      :      :

RESULT 5
Q98CR1 PRELIMINARY; PRT; 694 AA.
ID Q98CR1
AC Q98CR1
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein mlf5043.
GN Mlf5043.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,

```

```

RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL, AP003005; BAB51560.1; -.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004812; F:rRNA ligase activity; IEA.
DR GO; GO:0006418; P:amino acid activation; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR003781; CoA-binding.
DR InterPro; IPR003611; CoA_ligase.
DR InterPro; IPR001412; tRNA-synt_1.
DR Pfam; PF02629; CoA-binding; 1.
DR PROSITE; PS00178; AA TRNA LIGASE I; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 694 AA; 73126 MW; 19E4978531DCF6CC CRC64;

Query Match 16.4%; Score 82.5; DB 16; Length 694;
Best Local Similarity 33.3%; Pred.No. 25;
Matches 33; Conservative 11; Mismatches 42; Indels 13; Gaps 4;

QY 3 LAALLGLCVLSCSSAAAFVGS-AKPVAPVAALESAAAGAGTLANPLGLTNP-----L 57
:      :      :      :      :      :      :      :      :      :
Db 438 MVPLFGISEANDAGAAAFICGWAEEPAQPVDT--SAGGAGG-----GHVTPDEAEA 489
:      :      :      :      :      :      :      :      :      :

QY 58 KLLSSLGIPVNHLEGSQKVAELGQVAVGAVKALKAL 96
:      :      :      :      :      :      :      :      :      :
Db 490 KARLIKAGLPVPGKERAGNAVEAVISSMALGFPVALKAL 528
:      :      :      :      :      :      :      :      :      :

RESULT 6
Q9SU83 PRELIMINARY; PRT; 496 AA.
ID Q9SU83
AC Q9SU83;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nucleotide pyrophosphatase-like protein (EC 3.6.1.9).
GN TrfL4.190 OR ATG29680.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Rose M., Hempel S., Entian K.-D., Bancroft I., Mewes H.W.,
RA Mayer K.F.X., Lemcke K., Schueller C.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL079344; CAB45328.1; -.
DR EMBL; AL161575; CAB79726.1; -.
DR PIR; T09931; T09931.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004551; P:nucleotide diphosphatase activity; IEA.
DR GO; GO:0009117; P:nucleotide metabolism; IEA.
DR InterPro; IPR002591; Phosphodiester.
DR Pfam; PF01663; Phosphodiester; 1.

```

```

KW Hydrolase. 496 AA; 54678 MW; 8DC2B4346121D732 CRC64;
SQ SEQUENCE FROM N.A.
  Query Match 16.1%; Score 81; DB 10; Length 496;
  Best Local Similarity 41.3%; Pred. No. 24;
  Matches 26; Conservative 14; Mismatches 17; Indels 6; Gaps 3;

QY 6 LLGLCVLSCSSAAAEVLGSAKFPVQPVAAESAAEAG--GTLANPLGTIN-PLKLLLS 62
DB 57 LLVTCALSAASAFALFFSSQ---KPVLSLNQISKSPAFDSVARPLKLDKFPWLLIS 113

QY 63 SLG 65
DB 114 SDG 116

RESULT 7
ID Q8XW23 PRELIMINARY; PRT; 148 AA.
AC Q8XW23;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Probable lipoprotein.
GN RSC2652 OR RS04564.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21691879; PubMed=11923852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Catolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Iavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RA "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646071; CAD16359.1; -.
DR InterPro; IPR000437; Prok_lipoprot.s.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
RW Complete proteome.
SQ SEQUENCE 148 AA; 14704 MW; 7A5B624A84F58DB4 CRC64;

  Query Match 15.9%; Score 80; DB 16; Length 148;
  Best Local Similarity 31.1%; Pred. No. 8.1;
  Matches 37; Conservative 13; Mismatches 43; Indels 26; Gaps 7;

QY 3 LAALLGLCVLSCSSAAAEVLGSAKPVQA-----PVAALESAAEAG---AGTLANPL 51
DB 19 LAALGLCLLAACKSVPA--IDPKPPAMDPPAGNATPPASVSAANLGIIVG-LRAPT 75

QY 52 GTLNPLKLLSSLG-----IPVNHLEGSQKCVAEIG-PQAV--GAVKALKALIG 98
DB 76 DSAATLQTLNGICARATPPVAFVTPVPMGSGFWVRAASGSEPSATLDGAIALAAGP 134

RESULT 8
ID Q82R50 PRELIMINARY; PRT; 136 AA.
AC Q82R50;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative invasion protein.
GN SAV294.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RA "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=WA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RA "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AF005022; BAC68003.1; -.
KW Complete proteome.
SQ SEQUENCE 136 AA; 13687 MW; 8273376F70E3EBE7 CRC64;

  Query Match 15.7%; Score 79; DB 16; Length 136;
  Best Local Similarity 33.9%; Pred. No. 9.2;
  Matches 40; Conservative 11; Mismatches 43; Indels 24; Gaps 5;

QY 4 AALLGLCVLSCSSAAAEVLGSAKPVQAQPVAA-----LESAAEAGA 44
DB 20 STLMLITITLACLIALIFGAGVLTALPKARELAETGFSVAAYRRIGVLQAGVAGV 79

QY 45 --GTLANPLGTINPLKLLSSLGIPVNHLEGSQKCVAEIGPQAVGAVKALKALLGAL 100
DB 80 VLGLAVFPLGLAGAGLLGLLAGAVVHLROGDP--VAKLVPAAVCAV-LVASYLAL 134

RESULT 9
Q7WJ66 PRELIMINARY; PRT; 229 AA.
ID Q7WJ66
AC Q7WJ66;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative membrane protein.
GN B2634.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Basham S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Meule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RA "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640445; CA333127.1; -.
KW Complete proteome.
SQ SEQUENCE 229 AA; 22750 MW; 5A04498C4875195F CRC64;

  Query Match 15.4%; Score 77.5; DB 16; Length 229;
  Best Local Similarity 32.0%; Pred. No. 22;

```



Matches 32; Conservative 15; Mismatches 30; Indels 23; Gaps 4.

QY 11 VALSCSSAAAFVLSGAKPVAQPVAALESAAAGAGTLANPLGTINPLKILLSSLGIPVNH 70  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 144 VQAALASAASFAVGAAALPLAIAAAA-----FLAQLMPVVIAGSVAGLIGILG 189

QY 71 LIEGSQKCVAELGPQAV-----CAVKALKALIGALTVEG 104  
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :  
Db 190 AV-AAARGAPGVGAARVVLLGGAAAVALTAGVGL--FG 226

RESULT 10  
Q7WA34 PRELIMINARY; PRT; 229 AA.  
ID Q7WA34 AC Q7WA34  
DT 01-OCT-2003 (TREMBLrel. 25, Created)  
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Putative membrane protein.  
GN BPI1556  
OS Bordetella parapertussis.  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Alcaligenaceae; Bordetella.  
NCBI\_TaxID=519;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=12822 / ATCC BAA-587;  
RC MEDLINE=22827954; PubMed=12910271;  
RX Parkhill J., Sebaiha M., Preston A., Murphy L.D., Thomson N., Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L., Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A., Achman M., Atkin R., Baker S., Basham D., Basom N., Cherevach I., Challinorworth T., Collins M., Cronin A., Davis P., Doggett J., Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K., Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C., Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K., Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K., Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;  
RA "Comparative analysis of the genome sequences of Bordetella pertussis,  
RT Bordetella parapertussis and Bordetella bronchiseptica.";  
RL Nat. Genet. 35:32-40(2003).  
DR EMBL; EX640427; CAE36858.1; -.  
KW Complete proteome.  
SQ SEQUENCE 229 AA; 22750 MW; 5A04498C4875195F CRC64;

Query Match 15.4%; Score 77.5; DB 16; Length 229;  
Best Local Similarity 32.0%; Pred. No. 22;  
Matches 32; Conservative 15; Mismatches 30; Indels 23; Gaps 4;

QY 11 VALSCSSAAAFVLSGAKPVAQPVAALESAAAGAGTLANPLGTINPLKILLSSLGIPVNH 70  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 144 VQAALASAASFAVGAAALPLAIAAAA-----FLAQLMPVVIAGSVAGLIGILG 189

QY 71 LIEGSQKCVAELGPQAV-----CAVKALKALIGALTVEG 104  
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :  
Db 190 AV-AAARGAPGVGAARVVLLGGAAAVALTAGVGL--FG 226

RESULT 11  
Q7VVAL PRELIMINARY; PRT; 229 AA.  
ID Q7VVAL AC Q7VVAL  
DT 01-OCT-2003 (TREMBLrel. 25, Created)  
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Putative membrane protein.  
GN BPI1449  
OS Bordetella pertussis.  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Alcaligenaceae; Bordetella.  
NCBI\_TaxID=520;  
[1]  
RN SEQUENCE FROM N.A.

STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;  
MEDLINE=22827954; PubMed=12910271;  
Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.,  
Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,  
Achtman M., Atkin R., Baker S., Basham D., Bason N., Chervach I.,  
Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
Leather S., Moule S., Norberchak H., O'Neill S., Ormond D., Price C.,  
Rabinowitz E., Rutter S., Sanders M., Saunders D., Seeger K.,  
Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
Unwin L., Whitehead S., Barrall B.G., Maskell D.J.;  
"Comparative analysis of the genome sequences of *Bordetella pertussis*,  
*Bordetella parapertussis* and *Bordetella bronchiseptica*,"  
Nat. Genet. 35:32-40(2003).  
ENBL; BX640415; CAE41739.1; -.  
KW Complete proteome.  
SQ SEQUENCE 229 AA; 22776 MW; 5A1EF22D43DFA2FE CRC64;

Query Match 15.4%; Score 77.5; DB 16; Length 229;  
Best Local Similarity 32.0%; Pred. No. 22;  
Matches 32; Conservative 15; Mismatches 30; Indels 23; Gaps 4

QY 11 VALSCSSAAFLVGSAKPVAQPVFALESAAAGAGTIANPLGTINPLKLLSLGIPVNH 70  
Db 144 VQALASASPAVGAALPLATAMAA-----PLAQLMPVVIAGSVAGLIGL 189  
QY 71 LTGSGQKCVAGLGPQAV-----GAVKALKALLGNLTVFG 104  
Db 190 AV-AARAGAPVGPAAVRVVLGAAAMALTAGVGAL--FG 226

RESULT 12  
O86691 PRELIMINARY; PRT; 601 AA.  
ID AC O86691;  
AC O86691;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Putative transport system permease protein.  
GN SC06645 OR SC4G2.19  
OS Streptomyces coelicolor.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomycineae; Streptomycetaceae; Streptomyces.  
NCBI\_TaxId=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=B3(2) / M145;  
RX MEDLINE=21996410; PubMed=12000953;  
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieffer H.,  
Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
Huang C.-H., Kieffer T., Larke L., Murphy L., Oliver K., O'Neill S.,  
Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,  
Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,  
Hopwood D.A.;  
RT "Complete genome sequence of the model actinomycete *Streptomyces*  
coelicolor A3(2)," ;  
RL Nature 417:141-147(2002).  
DR ENBL; AJ939128; CAA20556.1; -.  
DR PIR; T35054; T35054.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR DR GO; GO:0006810; P:transport; IEA.  
DR DR InterPro; IPR000515; BPD transp.  
DR Pfam; PF00528; BPD transp. 2.  
KW Complete proteome.  
SQ SEQUENCE 601 AA; 61216 MW; 78601FCSAB2DCB9C CRC64;

Query Match 15.4%; Score 77.5; DB 16; Length 601;  
Best Local Similarity 30.7%; Pred. No. 63;

```

Matches 42; Conservative 10; Mismatches 48; Indels 37; Gaps 4;
Qy 3 LAALLGLCVALLSCSSAAAFVGS--KPVA-----QPVALESAAEAGAG----- 45
Db 426 LVALLVTAVAGSGAATPALAVGAVAWAPLAHTSSLLRQERATLHTATKGLGAGPVHLL 485
Qy 46 -----TLANPLGTINPLKLLSLGI-----PVNHLIEGSKCVKVAELGPQ 85
Db 486 RHELLFANVPVVRHALLRLPGVALALASLGLGLGAQPPSPENGLLLAENQPIAERAPW 545
Qy 86 AVGAKVAKKALIGALTIV 102
Db 546 AVLAPAAVALLGALAV 562

RESULT 13
Q8THLO PRELIMINARY; PRT; 453 AA.
AC Q8THLO;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein MA4504.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P., Brown A.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Maylor J., Stange-Thomann N., DeArellano K., Johnson R.,
RA Linton L., McKernan J., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuetznar H.C., Kzyzski J.A., Smith K.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., White R.H.,
RA Springer T.A., Umayal L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE011170; AA007844.1; -
DR GO; GO:0004590; P:orotidine-5'-phosphate decarboxylase activity; IEA.
DR GO; GO:0006207; P:'de novo' pyrimidine base biosynthesis; IEA.
DR InterPro; IPR005493; Methyltransf_6.
DR Pfam; PF03737; Methyltransf_6; 1.
DR Pfam; PF00215; OMPdecase; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 453 AA; 48593 MW; 4863027BB4AE65D CRC64;

Query Match 15.3%; Score 77; DB 17; Length 453;
Best Local Similarity 27.8%; Pred. No. 52;
Matches 30; Conservative 20; Mismatches 34; Indels 24; Gaps 4;
Qy 11 VALSCSSAAAFVGSAPVQAQPVVALESAAEAGAGTLANPLGTINPLK--LLLSGLI-- 66
Db 102 MAAKAGADVAIIIGSADD-STILDALRSAAHYGVVRVMDLISAPEPIKRAVDLBAIGVDY 160
Qy 67 -----PVNHLIEGSKCVKVAEL-----GPAQVGVAKVAKA 95
Db 161 INVHVGTDQVMGKDPISILMEISEKVSQVLAVAGSLDAEGRKAVKA 208

RESULT 14
Q8UA14 PRELIMINARY; PRT; 355 AA.
ID Q8UA14
AC Q8UA14;

```

```

DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE ABC transporter, membrane spanning protein.
GN ATU3389 OR AGR_L_2865.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Neeter E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmiel K., Gordon J., Vaudin M., Iatchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Garson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58";
RL Science 294:2323-2328(2001).
DR EMBL; AE009269; AAL44202.1; -
DR EMBL; AE008343; AAK90001.1; -
DR PIR; AD2973; AD2973.
DR PIR; G98309; G98309.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000522; FecD.
DR Pfam; PF01032; FecD; 1.
DR ProDom; PD001557; FecD; 1.
DR Complete proteome.
SQ SEQUENCE 355 AA; 37384 MW; 171A63B72224006C CRC64;

Query Match 15.1%; Score 76; DB 16; Length 355;
Best Local Similarity 23.3%; Pred. No. 49;
Matches 28; Conservative 22; Mismatches 42; Indels 28; Gaps 3;
Qy 2 KLAALGLCVALLSCSSAAAFVGSAPVQAQPVVALESAAEAGAGTLANPLGTINPLKLL 61
Db 24 RILVLGLFIALCFNSNAADMALGPARYTLSEVLA-----TIADPAAGVQLRVVI 73
Qy 62 SSLGIPVNHL-----IEGSKCVKVA-----ELGPAQVGVAKVAKKALIGALTIV 103
Db 74 WDIRMFIAMVIVGASLSVAGAQMQITLSNPLASPFITIGISAAASFGAALAVGVVAIF 133

RESULT 15
Q9ABLO PRELIMINARY; PRT; 699 AA.
ID Q9ABLO
AC Q9ABLO;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Thio:disulfide interchange protein, putative.
GN CC0217.
OS Caulobacter crescentus.

```

OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;  
 OC Caulobacteraceae; Caulobacter.  
 OX NCBI\_TaxID=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 19089 / CB15;  
 RX MEDLINE=21173698 / PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potočka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Ueberback T., Tran K., Wolf A., Vamathevan J., Emwolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of *Caulobacter crescentus*.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).  
 DR EMBL; AE005696; AAK22204.1; -  
 DR PIR; H87275; H87275.  
 DR TIGR; CC0217; -  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0007004; F:cycochrome biogenesis; IEA.  
 DR GO; GO:0006118; F:electron transport; IEA.  
 DR InterPro; IPR003834; Cytococh\_TM.  
 DR InterPro; IPR006663; ThioRedox\_dom2.  
 DR Pfam; PF02683; DsbD; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 699 AA; 70749 MW; 6523P9CA7E888943 CRC64;  
 Query Match 15.1%; Score 76; DB 16; Length 699;  
 Best Local Similarity 30.5%; Pred. NO. 1e+02;  
 Matches 32; Conservative 18; Mismatches 37; Indels 18; Gaps 5;  
 QY 3 LAALGLCVALSCTSAALFLVGSAPVAPVAALESAAEAGAGTLNPLGLTLNPLKLL 61  
 Db 545 VSMILGLUAVGALMAAASALSKP-----PVAAEASTPSGPGTAE---ANSPEKVAL 598  
 QY 62 SSLGIFVNHLEIGS-----QKCVAEIQPGVAVGAVKALKALL 97  
 Db 599 QAEGRPI--LVDFTAWCVTQVNEKVALSGPKVAEAFKAQNAV 641  
 RESULT 16  
 Q8G4X4  
 ID Q8G4X4 PRELIMINARY; PRT; 889 AA.  
 AC Q8G4X4;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Protease of ClpA/ClpB type.  
 GN CLPB OR BLI250.  
 OS Bifidobacterium longum.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;  
 OC Bifidobacteriaceae; Bifidobacterium.  
 OX NCBI\_TaxID=216816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCC 2705;  
 RX MEDLINE=22294977; PubMed=12381787;  
 RA Scheil M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,  
 RA Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,  
 RA Pridmore R.D., Arigoni F.;  
 RT "The genome sequence of *Bifidobacterium longum* reflects its adaptation  
 RT to the human gastrointestinal tract.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427 (2002).  
 DR EMBL; AB014751; AAN25051.1; -  
 DR GO; GO:0005224; F:ATP binding; IEA.  
 DR GO; GO:0003754; F:chaperone activity; IEA.  
 DR GO; GO:000166; F:nucleotide binding; IEA.  
 DR InterPro; IPR003593; AAA\_ATPass.  
 DR InterPro; IPR003593; AAA\_ATPass.  
 DR InterPro; IPR001270; Chaprinin\_clpA/B.  
 DR InterPro; IPR004176; Clp\_N.

Pfam; PF00004; AAA; 2.  
 Pfam; PF02861; Clp\_N; 2.  
 DR PRINTS; PRO0300; CLPPROTEASEA.  
 DR SMART; SMO0382; AAA; 2.  
 DR PROSITE; PS00870; CLPAB\_1; 1.  
 DR PROSITE; PS00871; CLPAB\_2; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 889 AA; 96313 MW; 704B8380EDA2A33C CRC64;  
 Query Match 15.1%; Score 76; DB 16; Length 889;  
 Best Local Similarity 25.7%; Pred. NO. 1.3e+02;  
 Matches 26; Conservative 21; Mismatches 36; Indels 18; Gaps 4;  
 QY 16 SSSAAFLVGSAPVAPVAALESAAEAGAGTLNPLGLTLNPL-----KILLS 63  
 Db 497 ABASKILYGIPIPSQKLAESAEDASADAANPAD--EPWVPRVDSVAEIVSDM 554  
 QY 64 LGIPVNHLEIGSQKVAE---LGPQAVGAVKALKALLGAL 100  
 Db 555 TGIPVGRIMQGENEKLHMDYLGKRVIGQKEAIAAVSDAV 595  
 RESULT 17  
 Q9F2L6  
 ID Q9F2L6 PRELIMINARY; PRT; 459 AA.  
 AC Q9F2L6;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative integral membrane protein.  
 GN SCO3765 OR SCH63.12C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Oliver K., Harris D.;  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96 (1996).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RX MEDLINE=21996410; PubMed=12000953;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,  
 RA Rabinovitch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 RT coelicolor A3(2)";  
 RL Nature 417:141-147 (2002).  
 DR EMBL; AL939117; CAC10303.1; -  
 DR InterPro; IPR002550; CBS.  
 DR InterPro; IPR000644; CBS\_domain.

```
DR InterPro; IPR005170; CorC_transp-asc.
DR Pfam; PF00571; CBS; 2.
DR Pfam; PF03471; CorC_HlyC; 1.
DR Pfam; PF01595; DUF21; 1.
DR SMART; SM00116; CBS; 1.
KW Complete proteome.
SQ SEQUENCE 459 AA; 48880 MW; E96CE64003F45752 CRC64;

Query Match          15.0%; Score 75.5; DB 16; Length 459;
Best Local Similarity 31.5%; Pred. No. 72;
Matches 34; Conservative 12; Mismatches 49; Indels 13; Gaps 5;

QY 3 LAALLGLCVALSCTSSAAFLVGSAPQVAALESAAEAGAGTTLNPLGTLNPKLLLS 62
DB 5 LLLLLALLLLTAC--ALFVAAFSLLTVERSLLERAAESGERGAGALRAVRSITLQLS 61
QY 63 --SIGIPVNHLEIG--SOKQVAIL--GPOAV----CAVKAALKALCAL 100
DB 62 GAGLIGTIVTSLVGMUAEPISVLLRGLPTAMGLGGAATVATLLGVV 109

RESULT 18
Q8G3S5 PRELIMINARY; PRT; 496 AA.
AC Q8G3S5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical secreted protein with D-Ala-D-Ala carboxypeptidase 3
DE (S13) domain.
GN BL1679.
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=NC 2705;
RX MEDLINE=2294577; PubMed=12381787;
RA Schell M.A., Kamirantzou M., Snel B., Villanova D., Berger B.,
RA Pessi G., Zwielen M.-C., Desiere F., Bork P., Delley M.,
RA Fridmore R.D., Arigoni F.;
RA "The genome sequence of Bifidobacterium longum reflects its adaptation
RA to the human gastrointestinal tract.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
DR EMBL; AB014802; AA025466.1; -.
DR GO; GO:0004185; F:serine carboxypeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR00667; Peptidase_S13.
DR Pfam; PF02113; Peptidase_S13; 1.
DR PRINTS; PR00922; DADACBPTASE3.
DR TIGRFams; TIGR00666; PRP4; 1.
KW Hypothetical protein; Carboxypeptidase; Complete proteome.
SQ SEQUENCE 496 AA; 50623 MW; 41572808ACD2D292 CRC64;

Query Match          15.0%; Score 75.5; DB 16; Length 496;
Best Local Similarity 30.3%; Pred. No. 78;
Matches 36; Conservative 15; Mismatches 43; Indels 25; Gaps 6;

QY 3 LAALLGLCVALSCTSSAAFLVGSAPQVAALESAAEAGAGTTLNPLGTLNPKL--- 59
DB 36 IALFAGYCAADITDVPGLL--TLKPVTAPVFADPATAKSG-GTVA---GTLNANKAIDS 89
QY 60 -----LLSLGI--PVNHLIEGSKQVA-----ELQFQAVGAVKALKALLGALT 102
DB 90 TAASALVNNLLSAQGVGNDSVIEDAQGVVAAEHESNTPEPASTLTKTLTALAASSTL 148

RESULT 19
Q9VEG2 PRELIMINARY; PRT; 544 AA.
ID Q9VEG2
AC Q9VEG2;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
```

```
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein AP20614.
GN AP20614.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococceae; Aeropyrum.
CX NCBI_TaxID=56636;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=K1.
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RA "Complete genome sequence of an aerobic hyper-thermophilic
RA crenarchaeon, Aeropyrum pernix K1.";
RT DNA Res. 6:83-101(1999).
RL EMBL; AP000060; BAA79584.1; -.
DR PIR; H72647; H72647.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005667; Lipocln_cytFABP.
DR PROSITE; PS00213; LIPCLNIN; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 544 AA; 55839 MW; 14028B4D558A7891 CRC64;

Query Match          15.0%; Score 75.5; DB 17; Length 544;
Best Local Similarity 26.4%; Pred. No. 86;
Matches 33; Conservative 19; Mismatches 36; Indels 37; Gaps 5;

QY 5 ALLGLCVALSCTSSA-----AAFVGSAPQVAA-----LESAEAGAGTLANP 50
DB 18 ALVAVAVAVAVAAALMSVGGRTAGLAGEAASQSIASAASNPVMEAVVAGA----- 71
QY 51 LGTLNPLKLLSLGIPVNHLLT-----EGSQKVAELGPOAVGAVKALK-----ALIGA 99
DB 72 -----LKRFFSPGAPVSVVVLRGEGVAAVRLPEFPAVEGVVDMEGYDCSPVLGV 125
QY 100 LTVFG 104
DB 126 ETVGG 130

RESULT 20
Q8XRM8 PRELIMINARY; PRT; 667 AA.
ID Q8XRM8
AC Q8XRM8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Probable signal peptide protein.
GN RSP083 OR RS01904.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OG Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=GMI1000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Ariat M., Billault A., Brottier P., Camus J.-C., Cattoi L.,
RA Chaudier M., Choise N., Claudel-Renard C., Cunnac S., Denange N.,
RA Gaspin C., Javie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
```

DR EMBL; AL646081; CAD17954.1; -  
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
KW Plasmid; Complete proteome.  
SQ SEQUENCE 667 AA; 59635 MW; 7C78268AC33B9E2B CRC64;  
  
Query Match 15.0%; Score 75.5; DB 16; Length 667;  
Best Local Similarity 29.1%; Pred. No. 1.1e+02;  
Matches 30; Conservative 14; Mismatches 44; Indels 15; Gaps 3;  
  
QY 14 SCSSAAAFVGSAXPQVPAQVVALESAAEAGAGTIANPLGTLNPKLLSGLG 65  
DB 498 SVSGSGALAPVQGVYVQVTLGSGNPTGALSGVGVGAVAGGSLGVSIG 557  
  
QY 66 IPVNH-----IRGSKQVAEIQPQVGAQVAVKALKALLGALT 101  
DB 558 APVSQVGSLSIPVGGSGVGVQVGAQVAVAGGSLGVS 600  
  
RESULT 21  
Q89XU3 PRELIMINARY; PRT; 299 AA.  
AC Q89XU3;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE B110213 protein.  
GN B110213.  
OS Bradyrhizobium japonicum.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Bradyrhizobium.  
OX NCBI\_TaxID=375;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=USDA 110;  
RX MEDLINE=22484998; PubMed=12597275;  
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,  
RA Sasamoto S., Watanabe A., Ideasa K., Iriguchi M., Kawashima K.,  
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,  
RA Tabata S.;  
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
Bradyrhizobium japonicum USDA110.";  
RL DNA Res. 9:189-197(2002).  
DR EMBL; AF005935; BAC45478.1; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR000620; DUF6.  
DR Pfam; PF00892; DUF6; 1.  
KW Complete proteome.  
SQ SEQUENCE 299 AA; 31937 MW; 86F2F929C7B18E94 CRC64;  
  
Query Match 14.9%; Score 75; DB 16; Length 299;  
Best Local Similarity 25.5%; Pred. No. 50;  
Matches 25; Conservative 22; Mismatches 47; Indels 4; Gaps 1;  
  
QY 3 LAALIGLCVSCSAAAFVGSAXPQVPAQVVALESAAEAGAGTIANPLGTLNPKLLS 62  
DB 81 LSGIVGTGLMLLAMDPSFVVTAYLKTETAI---QTAIFGVFLGDHLTLKVLAIWA 136  
  
QY 63 SIGIPVNHIEGSKQVAEIQPQVGAQVAVKALKALLGAL 100  
DB 137 TIGVWITALPQGGKSFALKEKTELIGLVAAAFALSAV 174  
  
RESULT 22  
Q8RPK5 PRELIMINARY; PRT; 355 AA.  
AC Q8RPK5;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Faga.  
GN Faga.  
OS Corynebacterium pseudotuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.  
OX NCBI\_TaxID=1719;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Whetten 1;  
RX MEDLINE=21932508; PubMed=11934492;  
RA Billington S.J., Emay P.A., Songer J.G., Jost B.H.;  
RT "Identification and role in virulence of putative iron acquisition  
genes from Corynebacterium pseudotuberculosis.";  
RL FEMS Microbiol. Lett. 208:41-45(2002).  
DR EMBL; AF401634; AAL79811.1; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005215; P:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000522; FecD.  
DR Pfam; PF01032; FecCD; 1.  
SQ SEQUENCE 355 AA; 36205 MW; F4DCF004C98FCBB CRC64;  
  
Query Match 14.9%; Score 75; DB 2; Length 355;  
Best Local Similarity 28.2%; Pred. No. 61;  
Matches 29; Conservative 24; Mismatches 32; Indels 18; Gaps 5;  
  
QY 12 ALSCSSAAAF-----LVGSAXPQVPAQVVALESAAEAGAGTIANPLG-----TLNPKLL 59  
DB 116 ALGINSGRAPFAIVVGIALFGASSPWFALALLGA--CGAAAVVSVGTHRSATADPVL 173  
  
QY 60 LLSGLIPVNHIEGSKQVAEIQPQVGAQVAVKALKALLGALT 102  
DB 174 VLS--GVALSAILSGIGELSLVNPQAFDRKSN--MVGSVDV 212  
  
RESULT 23  
Q89H46 PRELIMINARY; PRT; 370 AA.  
AC Q89H46;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE ABC transporter permease protein.  
GN BLR6145.  
OS Bradyrhizobium japonicum.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Bradyrhizobium.  
OX NCBI\_TaxID=375;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=USDA 110;  
RX MEDLINE=22484998; PubMed=12597275;  
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,  
RA Sasamoto S., Watanabe A., Ideasa K., Iriguchi M., Kawashima K.,  
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,  
RA Tabata S.;  
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
Bradyrhizobium japonicum USDA110.";  
RL DNA Res. 9:189-197(2002).  
DR EMBL; AF005957; BAC51414.1; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005215; P:transporter activity; IEA.  
DR InterPro; IPR001851; Bac\_inmem\_transp.  
DR Pfam; PF02653; BPD\_transp\_2; 1.  
KW Complete proteome.  
SQ SEQUENCE 370 AA; 613481F38466F22A CRC64;  
  
Query Match 14.9%; Score 75; DB 16; Length 370;  
Best Local Similarity 28.2%; Pred. No. 63;  
Matches 33; Conservative 17; Mismatches 45; Indels 22; Gaps 5;  
  
QY 3 LAALIGLCV-----ALSCSSAAAFVGSAXPQVPAQVVALESAAEAG-----AGTL 47  
DB 96 LPAQLGMVIGGEGALLIGALSATSALALQGMPLVQVIAWVIGGGLWMLSGAL 155  
  
QY 48 ANPLG---TLNPKLLSLGIPVNHIEGSKQVAEI---GPQAVGAVKALKALLG 98

156 ROYGVNETISSLLVYIALAI-LNELVEGLMFDPAISLKNKPTREIGANMIGSIPG 211

Db Q8C313 PRELIMINARY; PRT; 463 AA.

AC Q8C313; 23, Created)

DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)

DE SNFI-like kinase (fragment).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

EN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Lung;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

DR EMBL; AK087322; BAC39845.1; -.

FT NON-TER

SQ SEQUENCE 463 AA; 48966 MW; DA6327DB4AF27C78 CRC64;

Query Match 14.9%; Score 75; DB 11; Length 463;

Best Local Similarity 34.2%; Pred. No. 81;

Matches 27; Conservative 10; Mismatches 30; Indels 12; Gaps 3;

QY 10 CVALSCSAAFLVGSAPKPAQVPAALSAEAGAGTLANP--LGTLPKLLSLGIP 67

Db 174 CVIVSSGATSPSGTSDSCLPFSAGPAGLGS-GATPGLGTSVPRLASPL--- 229

QY 68 VNHLIEGSKCVABIGPQA 86

Db 230 -----GQSATPVLQTQA 242

RESULT 25

Q9L0K1 PRELIMINARY; PRT; 611 AA.

AC Q9L0K1; 15, Created)

DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)

DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)

DE Putative membrane protein.

GN SCO4664 OR SCD40A.10C.

OS Streptomyces coelicolor.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI\_TaxID=1902;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RA Seeger K.J., Harris D.;

RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RX MEDLINE=97000351; PubMed=8843436;

RA Redenbach M., Kiese H.M., Denapaita D., Eichner A., Cullum J.,

RA Kinashi H., Hopwood D.A.;

RT "A set of ordered cosmid, and a detailed genetic and physical map for

RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."

Mol. Microbiol. 21:77-96 (1996).

[4]

RL SEQUENCE FROM N.A.

RN STRAIN=A3(2) / M145;

RC MEDLINE=21996410; PubMed=12000953;

RX Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,

RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,

RA Harper D., Batenan A., Brown S., Chandra G., Chen C.W., Collins M.,

RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,

RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,

RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,

RA Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,

RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,

RA Hopwood D.A.;

RT "Complete genome sequence of the model actinomycete Streptomyces

RT coelicolor A3(2).";

RL Nature 417:141-147(2002).

DR EMBL; AL899121; CAB81855.1; -.

DR GO; GO:0016020; C:membrane; IEA.

DR InterPro; IPR003838; DUF214.

DR Pfam; PF02687; FtsX; 2.

KW Complete proteome.

SQ SEQUENCE 611 AA; 61744 MW; 8F959720773AEDF9 CRC64;

Query Match 14.7%; Score 74; DB 16; Length 611;

Best Local Similarity 41.9%; Pred. No. 1.3e+02;

Matches 26; Conservative 9; Mismatches 21; Indels 6; Gaps 2;

QY 3 LAALLGLCVALSAAFLVGSAPKPAQVPAALSAEAGAGTLANPGLTLPKLLIS 62

Db 134 LAFWTGLLVALCGAAASWAGRTGPVE---ALREAADAGATTGR---WLSGLALLT 187

QY 63 SL 64

Db 188 AL 189

RESULT 26

Q86DI1 PRELIMINARY; PRT; 119 AA.

AC Q86DI1;

DT 01-JUN-2003 (TREMELrel. 24, Created)

DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)

DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)

DE Stage-specific S antigen-like protein.

OS Leishmania infantum.

OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.

OX NCBI\_TaxID=5671;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MHOM/FR/92/LEM2385;

RA Garin Y.J.-F., Menseur P., Lorenzo F., Bui H., Pratlong F.,

RA Blanche H., Dedet J.-P., Derouin F.;

RA Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY255809; AAP21105.1; -.

SQ SEQUENCE 119 AA; 11523 MW; E32E2EC612D720EC CRC64;

Query Match 14.6%; Score 73.5; DB 5; Length 119;

Best Local Similarity 28.8%; Pred. No. 26;

Matches 30; Conservative 16; Mismatches 33; Indels 25; Gaps 4;

QY 1 MKLAALLGLCVALSAAFLVGSAPKPAQVPAALSAEAGAGTLANPGLTLPKLL 60

Db 1 MKIRSVRPLVLLVCVAAVLALGASAE-----HKAAYDAG-----PLSYD 41

QY 61 LSSLGIPVNHIEGSKCVABIGPQAVG--AVKALKALLGALT 102

Db 42 VGPLSVDPGLSVGPQ-----SVGPLSVGPQSDPLSVDPGLSV 81

RESULT 27

Q9HNI9 PRELIMINARY; PRT; 331 AA.

ID Q9HNI9

```

AC Q9HNI9,
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transport protein.
GN PHNE OR VNG2084G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shrona J.,
RA Swartzell S., Weir D., Hall J.A., Dahl T.A., Welci R., Goo Y.A.,
RA Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genomic sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AF005098; AAC20231.1; -.
DR F01; C94358; C84358.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPT_transp.
DR InterPro; IPR005479; CPase_L_D2.
DR Pfam; PF00528; BPT_transp; 1.
DR PROSITE; PS00867; CPASE_2; 1.
KW Complete proteome.
SQ SEQUENCE 331 AA; 36613 MW; 5AC660A5C6EAOE44 CRC64;

Query Match 14.6%; Score 73.5; DB 17; Length 331;
Best Local Similarity 27.9%; Pred. No. 77;
Matches 29; Conservative 17; Mismatches 35; Indels 23; Gaps 5;

QY 3 LAALGLCVSLSCSAAAFVGSAPQVPAALSAEAGAGTLANPLGTLPKLLLS 62
Db 83 LTSFGVVFVDVGQYGFSLAAHNP-SIPLAAVETGLIAPAGTV---LGA-PLALTFG 136
QY 63 SIG-----IPVNHIEGSKQVLAELGPQAVGAVKALKALIGAL 100
Db 137 VLGSERVTPFPLNFLFRG-----VMSIRSIPALVWAL 169

RESULT 28
Q8NQ28
ID Q8NQ28 PRELIMINARY; PRT; 350 AA.
AC Q8NQ28;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ABC-type transporter, permease components.
GN CGL1626.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005279; BAB99019.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001851; Bac_inmem_transp.

```

```

DR InterPro; IPR000522; FecD.
DR Pfam; PF01032; FecCD; 1.
DR ProDom; PD001557; FecCD; 1.
KW Complete proteome.
SQ SEQUENCE 358 AA; 37275 MW; 7F3AA9BF28BE2607 CRC64;

Query Match 14.6%; Score 73.5; DB 16; Length 358;
Best Local Similarity 26.5%; Pred. No. 84;
Matches 27; Conservative 19; Mismatches 45; Indels 11; Gaps 3;

QY 6 LLGLCVSLSCSAAAFVGSAPQVPAALSAEAGAGT-----LANPLGTLPK 58
Db 115 ILGVNSGASCGAAALLFGVGAGFGD--VALQSAFLGMAASGLIFFVARAAGRISSTR 172
QY 59 LLLSLGLIPVNHIEGSKQVLAELGPQAVGAVKALKALIGAL 100
Db 173 LIMS--GAIGYMLSAATSFLIFSSDSAGRSRVLFLLGSL 212

RESULT 29
Q8EJ01
ID Q8EJ01 PRELIMINARY; PRT; 389 AA.
AC Q8EJ01;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Prophage MuSol, protein Gp32, putative.
GN SO0674
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsipin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Ramathavan J., Weidman J., Imbraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Fiedlmyr T.V., Smith H.C., Venter J.C., Neilson K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015513; AAN53752.1; -.
DR TIGR; SO0674; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001337; TMV_coat.
KW Complete proteome.
SQ SEQUENCE 389 AA; 41160 MW; 42413806B5706ABB CRC64;

Query Match 14.6%; Score 73.5; DB 16; Length 389;
Matches 29; Conservative 12; Mismatches 26; Indels 21; Gaps 4;

QY 32 PVAALSAEAGAGTLANPLG-----TLNPLKL-LLSSLGIPVNHIEGSKQVLAEL 82
Db 173 PLALALAAALDTSINALPHTQPBGLHVNPLMKLLAALGI---ETAGSEPTAQL 229
QY 83 GP-----QAVGAVKALKALIGALT 101
Db 230 NDALKRIDDLVKAAGSVGDKQVEVAULT 257

RESULT 30
Q9GUB5
ID Q9GUB5 PRELIMINARY; PRT; 1468 AA.
AC Q9GUB5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)

```

QY 4 AAILG-LCVALSCSSAAFLVGS AKPVAALESAEAGAGTL-----ANPLGTL 54



OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PAOI;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Golltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Sailer M.H., Hancock R.B.W., Lory S., Olson M.V.,  
 RA "Complete genome sequence of Pseudomonas aeruginosa PAOI, an  
 RT opportunistic pathogen."  
 RL Nature 406:959-964(2000).  
 DR EMBL; AB004509; AAG04136.1; --  
 DR PIR; F83553; F83553.  
 DR HSP; P56533; 1A49.  
 DR GO; GO:0016491; P:oxidoreductase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR002086; Aldehyde\_dehydr.  
 DR Pfam; PF00171; aldehyd; 1.  
 DR Complete proteome.  
 SQ SEQUENCE 502 AA; 54564 MW; 33B2ABDFDA3B4973 CRC64;  
 Query Match 14.5%; Score 73; DB 16; Length 502;  
 Best Local Similarity 29.1%; Pred. No. 1.3e+02;  
 Matches 32; Conservative 13; Mismatches 41; Indels 24; Gaps 5;  
 QY 4 AALLGLCVLSCSSA-----AALVGSK-----PVAQPVAALES-----AAGAGTL 47  
 Db 264 AQLVGLNVLGSCGAAGRCMAISAAVFGAAREWIPELAERMAVLPGHMQDPDAAYGPL 323  
 QY 48 ANPLGLTNLPLKLLSLGIPVNHLEGSQKVAE-----LGPQAVGAV 90  
 Db 324 ISQARQVRILIAEGKAGACSLDGSQ-CQVEGYPNGNWLGTLPFAV 372  
 RESULT 34  
 O8GYAL PRELIMINARY; PRT; 542 AA.  
 AC O8GYAL  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Hypothetical protein (Atlg04570).  
 GN Atlg04570/tig11.17.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv, Columbia;  
 RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,  
 RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,  
 RA Hayashizaki Y., Shinozaki K.;  
 RA "Arabidopsis thaliana full-length cDNA."  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,  
 RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Huan V.W., Lam B.,  
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,  
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,  
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RA "Arabidopsis ORF clones."  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 RN EMBL; AK117774; BAC42421.1; --

DR EMBL; BT005916; AAO64851.1; --  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . . ; IEA.  
 DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
 DR InterPro; IPR004324; BT1.  
 DR InterPro; IPR001547; Glyco\_hydro\_5.  
 DR Pfam; PF03092; BT1; 1.  
 DR TIGRfams; TIGR00788; fbt; 1.  
 DR PROSITE; PS00659; GLYCOSYL\_HYDROL\_F5; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 542 AA; 59479 MW; 839B45F3FDD261B8 CRC64;  
 Query Match 14.5%; Score 73; DB 10; Length 542;  
 Best Local Similarity 28.2%; Pred. No. 1.5e+02;  
 Matches 35; Conservative 16; Mismatches 49; Indels 24; Gaps 4;  
 QY 3 LAALLGLCVLSCSSAAAFVGSAPVAPVAALSAEA-GAGTLANPLGLTNPLKLL 61  
 Db 233 LGSLLGGYLLTTPPKISFLVFSALLSLQLVLSKSESGFLPRTAETSSVLESVKQI 292  
 QY 62 SSL-----GIPVNHLEGSQK-----VAELGPQAVGAKAL-KALIG 98  
 Db 293 SNLKEAIQADEISQPLIWAUVVSIAMVLLSGSVFCYQTVLNLDPSSVIGMSKVGQMLL 352  
 QY 99 ALTV 102  
 Db 353 CLTV 356  
 RESULT 35  
 O9AAR5 PRELIMINARY; PRT; 584 AA.  
 ID O9AAR5  
 AC O9AAR5  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Gamma-glutamyltransferase.  
 GN CC0531  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;  
 OC Caulobacteraceae; Caulobacter.  
 OX NCBI\_TaxID=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 19089 / CB15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feidblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.B., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Usterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RA "Complete genome sequence of Caulobacter crescentus."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL; AB005725; AAK22518.1; --  
 DR PIR; B87315; B87315.  
 DR MEROPS; T03.001; --  
 DR TIGR; CC0531; --  
 DR GO; GO:0003840; P:gamma-glutamyltransferase activity; IEA.  
 DR GO; GO:0016740; P:transferase activity; IEA.  
 DR InterPro; IPR000101; Peptidase\_T3.  
 DR Pfam; PF01019; G\_glu\_transpept; 1.  
 DR PRINTS; PR01210; GGTTRANSPTASE.  
 DR TIGRfams; TIGR00056; g\_glu\_trans; 1.  
 KW Transferase; Complete proteome.  
 SQ SEQUENCE 584 AA; 51016 MW; 5C3DA03B02261E75 CRC64;  
 Query Match 14.5%; Score 73; DB 16; Length 584;  
 Best Local Similarity 28.6%; Pred. No. 1.6e+02;  
 Matches 28; Conservative 10; Mismatches 38; Indels 22; Gaps 2;  
 QY 2 KLAALLGLCVLSCSSAAAFVGSAPVAPV--AALESAAEAGAGTLANPLGLTNPLKL 59

DB	7	RLASLALSALQSLAPVAALAES:PLAWPTTPAATSTPAKGMVAANPLAVEAGURV	66
QY	60	LLSSLGIPVNHLEIGSKOAEVLELGPQAVGAVKALKALL	97
DB	67	LRD-----GGSAVDAVAIAQAVL	84
RESULT 36			
Q7WL37		PRELIMINARY;	PRT; 620 AA.
AC	Q7WL37;		
DT	01-OCT-2003 (TREMELrel. 25, Created)		
DT	01-OCT-2003 (TREMELrel. 25, Last sequence update)		
DT	01-OCT-2003 (TREMELrel. 25, Last annotation update)		
DE	1-deoxy-D-xylulose 5-phosphate synthase.		
GN	DXS OR B81912.		
OS	Bordetella pertussis (Alcaligenes bronchisepticus).		
OC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;		
OC	Alcaligenaceae; Bordetella.		
OK	NCBI_TaxID=518;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=RB50 / ATCC BAA-588;		
RC	MEDLINE=22827954; PubMed=12910271;		
RA	Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,		
RA	Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,		
RA	Achtman M., Atkin R., Temple L., James K., Harris B., Quail M.A.,		
RA	Cerdano-Tarraga A.M., Cronin A., Davis P., Doggett J.,		
RA	Chillingworth T., Collins M., Hauser N., Holroyd S., Jagels K.,		
RA	Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,		
RA	Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,		
RA	Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,		
RA	Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,		
RA	Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,		
RT	"Comparative analysis of the genome sequences of Bordetella pertussis,		
RT	Bordetella parapertussis and Bordetella bronchiseptica."		
RL	Nat. Genet. 35:32-40(2003).		
DR	EMBL; BX640442; CAE32409.1; -.		
KW	Complete proteome.		
SQ	SEQUENCE 620 AA; 66230 MW; 5D023E312C76D694 CRC64;		
Query Match	14.5%; Score 73; DB 16; Length 620;		
Best Local Similarity	31.1%; Pred. No. 1.7e+02;		
Matches	28; Conservative 11; Mismatches 31; Indels 20; Gaps 3;		
QY	12	ALSCSSAAAFVLSAKPVAQPVAALESAAEAGAGTLANPLGTLNPKLLSSLGIPVNH	71
DB	548	ALVTVEAAINGGAGS-----AVLETLAAG-----VTLFVLQGLGPDPAFI	588
QY	72	IEGSKC-VAEIGPQAVGAVKALKALLGAL	100
DB	589	DHGDQALLAGLGLDAAAGIERAIRARFGAL	618
RESULT 38			
Q7VV87		PRELIMINARY;	PRT; 620 AA.
ID	Q7VV87		
AC	Q7VV87;		
DT	01-OCT-2003 (TREMELrel. 25, Created)		
DT	01-OCT-2003 (TREMELrel. 25, Last sequence update)		
DT	01-OCT-2003 (TREMELrel. 25, Last annotation update)		
DE	1-deoxy-D-xylulose 5-phosphate synthase.		
GN	DXS OR B2738.		
OS	Bordetella pertussis.		
OC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;		
OC	Alcaligenaceae; Bordetella.		
OK	NCBI_TaxID=520;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;		
RC	MEDLINE=22827954; PubMed=12910271;		
RA	Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,		
RA	Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,		
RA	Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,		
RA	Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,		
RA	Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,		
RA	Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,		
RA	Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,		
RA	Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,		
RA	Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,		
RA	Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,		
RT	"Comparative analysis of the genome sequences of Bordetella pertussis,		
RT	Bordetella parapertussis and Bordetella bronchiseptica."		
RL	Nat. Genet. 35:32-40(2003).		
DR	EMBL; BX640442; CAE32409.1; -.		
KW	Complete proteome.		
SQ	SEQUENCE 620 AA; 66230 MW; 5D023E312C76D694 CRC64;		
Query Match	14.5%; Score 73; DB 16; Length 620;		
Best Local Similarity	31.1%; Pred. No. 1.7e+02;		
Matches	28; Conservative 11; Mismatches 31; Indels 20; Gaps 3;		
QY	12	ALSCSSAAAFVLSAKPVAQPVAALESAAEAGAGTLANPLGTLNPKLLSSLGIPVNH	71
DB	548	ALVTVEAAINGGAGS-----AVLETLAAG-----VTLFVLQGLGPDPAFI	588
QY	72	IEGSKC-VAEIGPQAVGAVKALKALLGAL	100
DB	589	DHGDQALLAGLGLDAAAGIERAIRARFGAL	618
RESULT 37			
Q7W700		PRELIMINARY;	PRT; 620 AA.
ID	Q7W700		
AC	Q7W700;		
DT	01-OCT-2003 (TREMELrel. 25, Created)		
DT	01-OCT-2003 (TREMELrel. 25, Last sequence update)		
DT	01-OCT-2003 (TREMELrel. 25, Last annotation update)		
DE	1-deoxy-D-xylulose 5-phosphate synthase.		
GN	DXS OR B27464.		
OS	Bordetella parapertussis.		
OC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;		
OC	Alcaligenaceae; Bordetella.		
OK	NCBI_TaxID=519;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=12822 / ATCC BAA-587;		
RC	MEDLINE=22827954; PubMed=12910271;		
RA	Parkhill J., Sebailia M., Preston A., Murphy L.D., Thomson N.,		

Db 589 DHGQDALLAGLGLDAAAGIERAIRARFGAL 618

## RESULT 39

O46099 PRELIMINARY; PRT; 1279 AA.

AC O46099; 1279 AA.

DT 01-JUN-1998 (TREMBlrel. 06, Created)

DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE EG:808.2 protein.

GN EG:808.2 OR CG11409.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]

RN SEQUENCE FROM N.A.

RC STRAIN=Berkely;

RC MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abriel J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,

RA Ballew R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cusley S., Dahlike C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,

RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K.A., Nusskern D.R., Pacle J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinart K., Renington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stappleton M., Strong R., Sun E.,

RA Svirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

RN [2]

RN SEQUENCE FROM N.A.

RA Papagiannakis G., Spanos L., Cox S., Siden-Kiamos I., Louis C.;

RT "Sequencing the distal X chromosome of Drosophila melanogaster.";

RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

RN [3]

RN SEQUENCE FROM N.A.

RA Benos P.;

RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE003420; AAP45604.1; -

DR EMBL; AJ022018; CAA17685.2; -

DR PIR; T13613; T13613.

DR FlyBase; FBgn0024366; EG:808.2.

SQ SEQUENCE 1279 AA; 139070 MW; FA96DBBA48C5D5E9 CRC64;

Query Match 14.5%; Score 73; DB 5; Length 1279;

Best Local Similarity 31.6%; Pred. NO. 3.7e+02;

Matches 31; Conservative 12; Mismatches 29; Indels 26; Gaps 4;

QY 25 SAKPVAQPVAALESA-----AEAGAGTLANPLGTLPKLLSLSLGIPVN 69

DB 213 SANPAAPSAARFSSHYSAKNAQFLRKPKSEGGGSLSS---TVKPVADILESGLIVSG 269

QY 70 HIEGSKQ-----CVAELGPQ-AGVAKKALKALLGA 99

DB 270 GKSDSAHKRYALDDYYPASAPQPSVAVADLRGLHGA 307

## RESULT 40

Q84CV5

ID Q84CV5

AC Q84CV5; PRELIMINARY; PRT; 240 AA.

DT 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Lipoprotein VacJ.

GN VACJ.

OS Gamma-proteobacterium Hot 75m4.

OG Plasmid pAK211.

OC Bacteria; environmental samples.

OX NCBI\_TaxID=77133;

RN [1]

RN SEQUENCE FROM N.A.

RC MEDLINE=22560762; PubMed=12673061;

RA Kniesch A., Waschkowitz T., Bowien S., Henne A., Daniel R.;

RT "Metagenomes of Complex Microbial Consortia Derived from Different

RT Soils as Sources for Novel Genes Confering Formation of Carbonyls

RT from Short-Chain Polyols on Escherichia coli.";

RL J. Mol. Microbiol. Biotech. 5:46-56(2003).

DR ENBL; AF548450; AAC91899.1; -

DR GO; GO:004821; C:extrachromosomal DNA; IEA.

DR InterPro; IPR000437; Prok\_lipoprot\_S.

DR InterPro; IPR007428; VacJ.

DR Pfam; PF04333; VacJ; 1.

DR PRINTS; PR01805; VACULIPOPROT.

DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; 1.

KW Lipoprotein; Plasmid

SQ SEQUENCE 240 AA; 25861 MW; FOC945D05D258E99 CRC64;

Query Match 14.4%; Score 72.5; DB 2; Length 240;

Best Local Similarity 27.8%; Pred. NO. 68;

Matches 35; Conservative 10; Mismatches 46; Indels 35; Gaps 4;

QY 7 LGLCVALSASSAAA-----FLVGSAPVAPVALESGAA- 40

DB 5 LPLIAALASLAAAGCTTTDPOSQAQNDPFBFSNRAINFIAVDKAVAPPAARFYRAAV 64

QY 41 -BAGAGTLANPLGTLPKLLSLSLGIPVNLIEG-SQKCVAEILGPQAVGAVKALKALLG 98

DB 65 PEPARQGVNLENLNAIPVL-----VNDVLOGEGEKAVNTGFRFVWNTVGLAGLID 117

QY 99 ALTVFG 104

DB 118 VATEFG 123

Search completed: April 5, 2004, 14:35:28

Job time : 43 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 5, 2004, 14:29:29 ; Search time 21 Seconds  
(without alignments)  
476.377 Million cell updates/sec

Title: US-09-997-428-408  
Perfect score: 502  
Sequence: 1 MKLAALLGLCVALSCHSAAA.....QAVGAVXALKALLGALTIVFG 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191536 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database :

PIR\_79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	16.1	496	2 T09931	probable phosphodi
2	79.5	15.8	1381	2 S60004	hypothetical prote
3	77.5	15.4	601	2 T35054	probable transport
4	77	15.3	335	2 T36304	probable anthranil
5	76	15.1	355	2 A26973	hypothetical prote
6	76	15.1	355	2 G98309	probable ABC trans
7	76	15.1	699	2 H87275	thio-disulfide int
8	76	15.1	732	2 T47269	copper-transportin
9	75.5	15.0	544	2 H72647	hypothetical prote
10	73.5	14.6	331	2 C84358	transport protein
11	73	14.5	502	2 F83553	probable aldehyde
12	73	14.5	584	2 B87315	gamma-glutamyltran
13	73	14.5	1279	2 T13610	hypothetical prote
14	72	14.3	236	2 A56010	anastigote-specifi
15	72	14.3	397	2 A52900	succinyl-CoA synth
16	72	14.3	397	2 F97675	succinyl-CoA synth
17	72	14.3	528	2 D70968	hypothetical prote
18	71.5	14.2	462	2 B87634	L-serine dehydrata
19	71.5	14.2	874	2 A30070	ATP-dependent Clp
20	71.5	14.2	887	2 F98216	endopeptidase clp
21	71	14.1	440	2 C83368	probable MFS trans
22	70.5	14.0	668	2 G85160	heat shock protein
23	70.5	14.0	831	2 D71409	probable endopepti
24	70	13.9	431	2 C86178	hypothetical prote
25	70	13.9	452	2 T46147	zinc finger protei
26	70	13.9	477	2 D82179	probable multidrug
27	69.5	13.8	244	2 S75653	hypothetical prote
28	69.5	13.8	396	2 G75454	hypothetical prote
29	69.5	13.8	423	2 C70582	probable PPE prote

probable L-serine  
hypothetical prote  
serine proteinase,  
probable ATP-depen  
probable adhesin Y  
hypothetical prote  
ferric enterobacti  
P-glycoprotein - S  
ferric enterobacti  
ferric enterobacti  
ferric enterobacti  
hypothetical prote  
branched-chain ami  
hypothetical prote  
probable conserved  
similar to transla  
glyoxylase II fami  
aspartate racemase  
probable glycoprot  
probable glycoprot  
leucine-specific b  
hypothetical prote  
cation-transportin  
hypothetical prote  
conserved hypothet  
CT483 hypothetical  
conserved hypothet  
hypothetical prote  
zein protein - mai  
hypothetical prote  
thromboxane A2 rec  
FREAC-4 - human  
hypothetical prote  
ecdysone-induced p  
probable minor tai  
lipid transfer pro  
lysr-family trans  
thiamin ABC transp  
hypothetical prote  
acetolactate synth  
hypothetical prote  
conserved hypothet  
hypothetical prote  
sspB protein - Sal  
F28N24.25 protein  
citrate synthase I  
hypothetical prote  
slpB protein - Sal  
pathogenicity isla  
heat shock protein  
ATP-dependent Clp  
hypothetical prote  
hypothetical prote  
hypothetical prote  
iron(III)-binding  
probable PPE prote  
alpha-2C-adrenergi  
DNA recombinase (E  
hypothetical prote  
ankyrin 3, long sp  
membrane-bound cyt  
hypothetical prote  
iron compound ABC  
hypothetical prote  
multidrug-efflux t  
hypothetical prote  
endoglycosylcerami  
conserved hypothet  
elastin precursor

ALIGNMENTS

RESULT 1  
T09931  
Probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) T16L4  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Aug-2002  
C/Accession: T09931  
R/Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.  
submitted to the Protein Sequence Database, June 1999  
A/Reference number: Z16897  
A/Accession: T09931  
A/Molecule type: DNA  
A/Residues: 1-496 <BEV>  
A/Cross-references: EMBL:AL079344; GSPDB:GN00062; ATSP:T16L4.190  
A/Experimental source: cultivar Columbia; EAC clone T16L4  
C/Genetics:  
A/Map position: 4  
C/Superfamily: human phosphodiesterase I / nucleotide pyrophosphatase 4  
C/Keywords: coenzyme A; metalloprotein; phosphoric diester hydrolase  
  
Query Match 16.1%; Score 81; DB 2; Length 496;  
Best Local Similarity 41.3%; Pred. No. 5.2;  
Matches 26; Conservative 14; Mismatches 17; Indels 6; Gaps 3;  
  
QY 6 LGLCVALSAAAPLVGSAKPVAPVAALSAEAGA--GTLANPLGTLN-PLKILLS 62  
DB 57 LLVTCIALSAASAFALFFSSQ---KPVLSLNGISKSPAFDRSVARPLKLDKPVVLLIS 113  
  
QY 63 SLG 65  
DB 114 SDG 116  
  
RESULT 2  
S60004  
Hypothetical protein - common roundworm retrotransposon R4 (fragment)  
C/Species: Ascaris lumbricoide (common roundworm)  
C/Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 21-Jul-2000  
C/Accession: S60004  
R/Burke, W.D.; Mueller, F.; Eickbush, T.H.  
Nucleic Acids Res. 23, 4628-4634, 1995  
A/Title: R4, a non-LTR retrotransposon specific to the large subunit rRNA genes of nematodes  
A/Reference number: S60004; MUID:96103592; PMID:8524653  
A/Accession: S60004  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-1381 <BUR>  
A/Cross-references: EMBL:U29445; NID:g903660; PIDN:AAA97394.1; PID:g903661  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995  
C/Genetics:  
A/Genome: retrotransposon  
  
Query Match 15.8%; Score 79.5; DB 2; Length 1381;  
Best Local Similarity 27.8%; Pred. No. 20;  
Matches 25; Conservative 15; Mismatches 23; Indels 27; Gaps 3;  
  
QY 11 VALSCSSAAFLVGSAPVAPVAALSAEAGAAGT-----LANPLG----- 52  
DB 27 IAMPCTNSFFERTGPFPHRSPRSLGTHRSPLNDDEVINGPKGHSDPVH 86  
  
QY 53 -----TLNPLKLLSSLGIPVNHIEGSQ 76  
DB 87 VVRAPRLHPRL---ELPIGVNVLGEASQ 113  
  
RESULT 3  
T35054  
Probable transport system permease protein - Streptomyces coelicolor  
C/Species: Streptomyces coelicolor  
C/Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999

C/Accession: T35054  
R/Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, August 1998  
A/Reference number: Z21566  
A/Accession: T35054  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-601 <SEE>  
A/Cross-references: EMBL:AL031371; PIDN:CAA20556.1; GSPDB:GN00070; SCOEDB:SC4G2.19  
A/Experimental source: strain A3(2)  
C/Genetics:  
A/Gene: SCOEDB:SC4G2.19  
  
Query Match 15.4%; Score 77.5; DB 2; Length 601;  
Best Local Similarity 30.7%; Pred. No. 13;  
Matches 42; Conservative 10; Mismatches 48; Indels 37; Gaps 4;  
  
QY 3 LAALGLCVALSAAAFVGSAA--KPVA-----QPVAALSAEAGAAG----- 45  
DB 426 LVALLVTAVAGSGAATPALAVGAVAWAPLAHTSLLRQERATLHITATKGLGAGPVHLL 485  
  
QY 46 -----TIANPLGTLNPLKLLSSLG-----PVNHLIEGSKQCVAEILGPQ 85  
DB 486 RHELLPAVVPVRLHALLRPGVALASLGLGLGAQPPSPFWGLLAENQPYAERAPW 545  
  
QY 86 AYCAVKALKALIGALTIV 102  
DB 546 AVLAAPAAVLALIGALAV 562  
  
RESULT 4  
T36304  
Probable anthranilate phosphoribosyltransferase - Streptomyces coelicolor  
C/Species: Streptomyces coelicolor  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Jan-2000  
C/Accession: T36304  
R/Saunders, D.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, March 1999  
A/Reference number: Z21604  
A/Accession: T36304  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-335 <SAU>  
A/Cross-references: EMBL:AL035654; PIDN:CAB38583.1; GSPDB:GN00070; SCOEDB:SC68.05C  
A/Experimental source: strain A3(2)  
C/Genetics:  
A/Gene: trpD2; SCOEDB:SC68.05C  
C/Superfamily: anthranilate phosphoribosyltransferase; trpD homology  
  
Query Match 15.3%; Score 77; DB 2; Length 335;  
Best Local Similarity 34.5%; Pred. No. 8.2;  
Matches 30; Conservative 9; Mismatches 36; Indels 12; Gaps 4;  
  
QY 7 LGLCVALSAAAFVGSAA--KPVAQPVAALSAEAA-----EAGAGTLANPLGTL-NP 56  
DB 117 LGVRIDLGAEAAACLDRTGITFLFAPVPH--APRHTAGPRRELGAITVFLGLPLCNP 174  
  
QY 57 LKLLSSLGIPVNHIEGSKQCVAEILG 83  
DB 175 SGARLRTLGVPSELVPEPMTVEILRG 201  
  
RESULT 5  
AD2973  
Hypothetical protein Atu3389 [imported] - Agrobacterium tumefaciens (strain C58, Dupont  
C/Species: Agrobacterium tumefaciens  
C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
C/Accession: AD2973  
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl1el  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

[illegible]

A:Molecule type: DNA  
A:Residues: 1-699 <STO>  
A:Cross-references: GB:AE005673; NID:g13421344; PIDN:AAX22204.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC0217

Query Match            15.1%; Score 76; DB 2; Length 699;  
Best Local Similarity   30.5%; Pred.No. 21;  
Matches     32; Conservative   16; Mismatches   37; Indels   18; Gaps   5;

QY      3 LAALLGLCVALSCSSAAAFVLVGSAPKVAQPVAALSAEAGAGATTIANPLGTLNPULKL-LL 61  
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db      545 VSMILGLLVAVGALAALASAKP---PVAAAEEASTPSGPLTAE----AWSPEKVQAL 598

QY      62 SSLGIPIPNHLEGS-----QKCVAELGPQAVGVAKALKALI 97  
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db      599 QAEGRFI--LVDFTAAMCVTCQNVNEKVALSGPKVAFAFQAQNAVLF 641

RESULT 8  
T47269  
Copper-transporing ATPase [EC 3.6.1.-] P-type copA [validated] - Helicobacter felis  
C:Species: Helicobacter felis  
C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 08-Sep-2000  
C:Accession: T47269  
R:Bayle, D.; Wangler, S.; Weitzenegger, T.; Steinhilber, W.; Volz, J.; Przybylski, M.; J. Bacteriol. 180, 317-329, 1998  
A:title: Properties of the P-type ATPases encoded by the copA operons of Helicobacter  
A:Reference number: Z44437; MUID:98101471; PMID:9440521  
A:Accession: T47269  
A>Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-732 <BAV>  
A:Cross-references: EMBL:AJ001932; NID:g2660538; PIDN:CAA05104.1; PID:g2660542  
A:Experimental source: strain ATCC 49179  
C:Genetics:  
A:Gene: copA  
C:Function:  
A>Description: binds copper specifically [validated, MUID:98101471]  
A>Note: involved in copper resistance  
C:Superfamily: Enterococcus copper-transporting ATPase copA; ATPase nucleotide-binding  
C:Keywords: copper binding; hydrolase  
F:8-37/Domain: heavy-metal-associated homology <HMA>  
F:532-674/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match            15.1%; Score 76; DB 2; Length 732;  
Best Local Similarity   26.3%; Pred.No. 22;  
Matches     35; Conservative   17; Mismatches   39; Indels   42; Gaps   6;

QY      6 LLGLCVALSCSSAAAFVLVGSAPKVAQPVAAL-----ESAAEAAGAG----- 45  
      LTTLCASLEAQSEHVIAKGIVAHAKEGIALQEVEQVKPGFGIKGWGDQIIKAGNUIE 506  
Db      447 LTTLCASLEAQSEHVIAKGIVAHAKEGIALQEVEQVKPGFGIKGWGDQIIKAGNUIE 506

QY      46 --TLANPLGTINPLKLLSS---LIGIPV--NHILEGSQKCAVELGPQAVGA----- 89  
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db      507 FNLPMFPFTLEGIOVFVGTEQTQLGVVVLDLSKEGSKAISSEL--XALGVKTILLSGD 564

QY      90 ----YXALKALLG 98  
      |:|||:|||||  
Db      565 NLENVZALATQLG 577

RESULT 9  
H72647  
hypothetical protein APE0614 - Aeropyrum pernix (strain Kl)  
C:Species: Aeropyrum pernix  
C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999  
C:Accession: H72647  
F:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; DNA Res. 6, 83-101, 1999  
A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix  
A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: H72647	
A:Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-544 <KAW>	
A:Cross-references: DBU:AF000060; NID:95104188; PIDN:BAA79584.1; PID:d1043370; PID:g510	
A:Experimental source: strain K1	
C:Genetics:	
A:Gene: AP80614	
Query Match 15.0%; Score 75.5; DB 2; Length 544;	
Best Local Similarity 26.4%; Pred. No. 18;	
Matches 33; Conservative 19; Mismatches 36; Indels 37; Gaps 5;	
QY	5 ALIGLGVALSCESSA-----RAFLVGSAPVAPVAA-----LESAAEAGAGTLANP 50
DB	18 ALVAVAVAVAAALMSVGRGTAGIAGAGAAASQIAASNPPLYMEANVEAGA----- 71
QY	51 LGTLNPLKLLSSLGIPVNHLL-----EGSKQKVAELGPOAVGAVALK-----ALIGA 99
DB	72 -----LKVRFYSPGAPVESVVLKPGEGVAARVELPEPAVEGVVDAMEGYPVILGV 125
QY	100 LTVFG 104
DB	126 ETVGG 130
RESULT 10	
C84358	
transport protein [imported] - Halobacterium sp. NRC-1	
C:Species: Halobacterium sp. NRC-1	
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001	
C:Accession: C84358	
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.	
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo	
Jung, K.H.; Alam, M.; Freitas, T.	
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000	
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li	
A:Title: Genome sequence of Halobacterium species NRC-1.	
A:Reference number: A84160; MUID:20504483; PMID:11016950	
A:Accession: C84358	
A:Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-331 <STO>	
A:Cross-references: GB:AE004437; NID:g10581500; PIDN:AAG20231.1; GSPDB:GN00138	
C:Genetics:	
A:Gene: phnE	
Query Match 14.6%; Score 73.5; DB 2; Length 331;	
Best Local Similarity 27.9%; Pred. No. 17;	
Matches 29; Conservative 17; Mismatches 35; Indels 23; Gaps 5;	
QY	3 LAALIGLGVALSCESSAAAFVGSAPVAPVAALESAAEAGAGTLANPLGTLNPLKLLS 62
DB	83 LTSLFQWVFVVDVGGYSGFLAAHENP-SIPLAAVETIGIAFAGTV---LGA--FLALTFG 136
QY	63 SLG-----IPVNHLEGSQKCVABLGPOAVGAVKALKALGAL 100
DB	137 VLGSERVTFPLNPLFRG-----VMSIRSIPALVWAL 169
RESULT 11	
F83553	
probable aldehyde dehydrogenase PA0747 [imported] - Pseudomonas aeruginosa (strain PA01)	
C:Species: Pseudomonas aeruginosa	
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000	
C:Accession: F83553	
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B	
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,	
; Lory, S.; Olson, M.V.	
Nature 406, 959-964, 2000	
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho	
A:Reference number: A82950; MUID:20437337; PMID:10984043	
A:Accession: F83553	
A:Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-502 <STO>	
A:Cross-references: GB:AE004509; GB:AE004091; NID:g9946622; PIDN:AAG04136.1; GSPDB:GN00	
A:Experimental source: strain PA01	
C:Genetics:	
A:Gene: PA0747	
C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology	
Query Match 14.5%; Score 73; DB 2; Length 502;	
Best Local Similarity 29.1%; Pred. No. 29;	
Matches 32; Conservative 13; Mismatches 41; Indels 24; Gaps 5;	
QY	4 AALLIGLGVALSCESSA-----AAFLVGSAX-----PVAQPVALES-----AERAGTLL 47
DB	264 AQVLGNLVGASCAAGRCWAISSAAVFGAAKEWIPELAEKRAVLRPGHWQDDPDAYGFL 323
QY	48 ANPLGTLNPLKLLSSLGIPVNHLEGSQKCVAE-----LGPOAVGAV 90
DB	324 ISFQARQVRLRIAESKAEAGACLLDGSQ-CQVEGYPNGWLGPTLFRAY 372
RESULT 12	
B87315	
gamma-glutamyltransferase [imported] - Caulobacter crescentus	
C:Species: Caulobacter crescentus	
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001	
C:Accession: 887315	
R:Niferman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J	
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Koic	
n, J.; Ermolaeva, M.; Waite, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M	
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001	
A:Title: Complete Genome Sequence of Caulobacter crescentus.	
A:Reference number: A87249; MUID:21173698; PMID:11259647	
A:Accession: 887315	
A:Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-584 <STO>	
A:Cross-references: GB:AE005673; NID:g13421716; PIDN:AAK22518.1; GSPDB:GN00148	
C:Genetics:	
A:Gene: CC0531	
C:Superfamily: gamma-glutamyltransferase	
Query Match 14.5%; Score 73; DB 2; Length 584;	
Best Local Similarity 28.6%; Pred. No. 34;	
Matches 28; Conservative 10; Mismatches 38; Indels 22; Gaps 2;	
QY	2 KLAALIGLGVALSCESSAAAFVGSAPVAPV--AALESAAEAGAGTLANPLGTLNPLKL 59
DB	7 RLASLLSAAALQUSLAPVAALESIPLAMPTRPRAATSTPAKGMVAANFLAVEAGLRV 66
QY	60 LLSSLGIPVNHLEGSQKCVAEELGPOAVGAVKALKALL 97
DB	67 LRD-----GGSAVDAAVAIAQAVL 84
RESULT 13	
T13613	
hypothetical protein 8D8.2 - fruit fly (Drosophila melanogaster)	
C:Species: Drosophila melanogaster	
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000	
C:Accession: T13613	
R:Papagiannakis, G.; Spance, L.; Cox, S.; Siden-Kiamos, I.; Louis, C.	
submitted to the EMBL Data Library, April 1999	
A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.	
A:Reference number: Z17695	
A:Accession: T13613	
A:Status: preliminary; translated from GB/EMBL/DBJ	
A:Molecule type: DNA	
A:Residues: 1-1279 <PAP>	
A:Cross-references: EMBL:AL022018; NID:e1273253; PID:e1426350; PIDN:CAAL7685.2	
C:Genetics:	
A:Cross-references: FlyBase:FBgn0024367	

```
A;Introns: 65/2
A;Note: EG:8DB.2

Query Match      14.5%; Score 73; DB 2; Length 1279;
Best Local Similarity 31.6%; Pred. No. 75;
Matches 31; Conservative 12; Mismatches 29; Indels 26; Gaps 4;

QY 25 SAKPVAQPVAALISA-----ABAGAGTLANPLGTLNPLKLLSLGIPVNVNLISSGQ 69
DB 213 SANPAAPSAARFSSHYSAKNAQFURKPKPSSGGGSLSS---TVKPVADILESLGIVSG 269
QY 70 HLIEGSK-----CVAELGPQ-AVGAVKALKALIGA 99
DB 270 GKSDSHKRYALDDYYPASAPQPSVAVADLRGLHGA 307

RESULT 14
A56010
amastigote-specific protein A2 precursor - Leishmania donovani infantum
C;Species: Leishmania donovani infantum
C;Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 24-Nov-1999
C;Accession: A56010
R;Charest, H.; Matlashewski, G. 1994
Mol. Cell. Biol. 14, 2975-2984, 1994
A;Title: Developmental gene expression in Leishmania donovani: differential cloning and
A;Reference number: A56010; MUID:94217695; PMID:7545921
A;Accession: A56010
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-236 <CHA>
A;Cross-references: GB:S69693; NID:9546453; PID:9546454
C;Superfamily: proline-rich protein
C;Keywords: tandem repeat

Query Match      14.3%; Score 72; DB 2; Length 236;
Best Local Similarity 29.8%; Pred. No. 17;
Matches 31; Conservative 15; Mismatches 42; Indels 16; Gaps 4;

QY 1 MKLAALLGHCVALSCSSAAAFVLSGAKP--VAQPVAALISAAGAGTLANPLGTLNPLK 58
DB 1 MKIRSVRLVLLVCAVAVLALSASAEHPKAAVDVGPL-SVGPQSVGLSVGPQAVGPLS 59
QY 59 LLLSLGIPVNHLEIAGSKVAVELGPQAVGAVKALKALGALT 102
DB 60 VGPQSVG-PLS-----VGPQAVGPLSVGPQSVGLSV 90

RESULT 15
AE2900
succinyl-CoA synthetase beta chain [imported] - Agrobacterium tumefaciens (strain C58, D
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C;Accession: AE2900
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.;
Erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan,
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AE2577; MUID:21608850; PMID:11743193
A;Accession: AE2900
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-397 <KUR>
A;Cross-references: GB:AE008688; PID:917741140; GSPDB:GN00186
A;Experimental source: strain C58 (Dupont)
C;Genetics:

A;Gene: succ
A;Map position: circular chromosome
C;Superfamily: succinate-CoA ligase (ADP-forming) beta chain

Query Match      14.3%; Score 72; DB 2; Length 397;

Best Local Similarity 36.2%; Pred. No. 28;
Matches 29; Conservative 8; Mismatches 31; Indels 12; Gaps 4;

QY 18 AAFLVGSAPVAPVAALISAAGAGTLANPLGTLNPLKLLSLGIPVNVNLISSGQ 76
DB 8 AKALLKGYPVAGVAILKVEEAEAAAKQLPGPLYV---VKSQIHAGG-----RGKG 57
QY 77 KCVAVELGPQAVGAVKALKAL 96
DB 58 K-FKELGPDPAKGGVRLAKSI 76

RESULT 16
F97675
succinyl-CoA synthetase beta chain (AF326913) [imported] - Agrobacterium tumefaciens (
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C;Accession: F97675
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldma
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium t
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: F97675
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-397 <KUR>
A;Cross-references: GB:AE007869; PID:AAK88359.1; PID:g15157842; GSPDB:GN00169
A;Genetics:
A;Gene: AGR_C_4780
A;Map position: circular chromosome
C;Superfamily: succinate-CoA ligase (ADP-forming) beta chain

Query Match      14.3%; Score 72; DB 2; Length 397;
Best Local Similarity 36.2%; Pred. No. 28;
Matches 29; Conservative 8; Mismatches 31; Indels 12; Gaps 4;

QY 18 AAFLVGSAPVAPVAALISAAGAGTLANPLGTLNPLKLLSLGIPVNVNLISSGQ 76
DB 8 AKALLKGYPVAGVAILKVEEAEAAAKQLPGPLYV---VKSQIHAGG-----RGKG 57
QY 77 KCVAVELGPQAVGAVKALKAL 96
DB 58 K-FKELGPDPAKGGVRLAKSI 76

RESULT 17
D70968
hypothetical protein RV2672 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: D70968
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genom
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70968
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-528 <COL>
A;Cross-references: GB:Z80225; GB:AL121456; NID:g3242265; PID:CAE02326.1; PID:e266414
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: RV2672

Query Match      14.3%; Score 72; DB 2; Length 528;
Best Local Similarity 28.4%; Pred. No. 38;
Matches 29; Conservative 15; Mismatches 34; Indels 24; Gaps 4;

QY 3 LAALLGCVALLSCSSAAAFVLSGAPVAPVAALISAAGAG-----TLANPLGTLN 55
```



```
Db 300 LDAFAACVAVNCA-----LGSHPKGA--VSALLSAARSGDGPQGSAAVAVNAVAT-- 349
Qy 56 PLKLLSSLGIPVNHLEGSQKCVAEIGPQAVGAVKALL 97
Db 350 -----ALGFPDSGRVDSITTKLADALAAARSGDMNLLSALI 384

RESULT 18
B87634
L-serine dehydratase [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: B87634
R:NIerman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173699; PMID:11259647
A:Accession: B87634
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-462 <STO>
A:Cross-references: GB:AE005673; NID:gl3424766; PIDN:AAK25070.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC3108
C:Superfamily: microbial L-serine dehydratase

Query Match 14.2%; Score 71.5; DB 2; Length 462;
Best Local Similarity 31.2%; Pred. No. 37;
Matches 30; Conservative 12; Mismatches 21; Indels 33; Gaps 6;
Qy 11 VALSCSAAAFVGSAPVQAPVAALESAAEAG-----AGTLANPLGTNPLKLLSSLGI 66
Db 347 VGVACSMAAA-----GLAALGGTNAQENAAEIGMEHNLGTCDPDGL-----VOI 394
Qy 67 PVNHLIEGSKCVAEIGPQAVGAVKALK-----ALLG 98
Db 395 P-----CTER---NANGAIKIDAARLALLG 417

RESULT 19
AC3070
ATP-dependent Clp proteinase, ATP-binding subunit clpB [imported] - Agrobacterium tumefa
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C:Accession: AC3070
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenthrner, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AC3070
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-874 <KUR>
A:Cross-references: GB:AE008689; PIDN:AAL44977.1; PID:gl7742634; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: clpB
A:Map position: linear chromosome
C:Superfamily: endopeptidase Clp ATP-binding chain

Query Match 14.2%; Score 71.5; DB 2; Length 874;
Best Local Similarity 26.1%; Pred. No. 70;
Matches 23; Conservative 23; Mismatches 37; Indels 5; Gaps 2;
Qy 18 AAAPVLSGAKVAPVQAVPVALESAAEAGAGTLANPLGTNPLKLLSS-LGIPVNHLEGSQ 76

Db 499 AGELTYGIIPGLEKELEAAAEARDSSGAGSMVQEVVTPDNIAHVVSRTWTGIPVDKMLEGQR 558
Qy 77 KCVA-----ELGPPQAVGAVKALKALLGAL 100
Db 559 EKLLRMEDELAKSVVGGQGEAVQAVSKAV 586

RESULT 20
F98216
endopeptidase clp ATP-binding chain B [imported] - Agrobacterium tumefaciens (strain C5
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
C:Accession: F98216
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: F98216
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-587 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK99256.1; PID:gl5159084; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L1346
A:Map position: linear chromosome
C:Superfamily: endopeptidase Clp ATP-binding chain

Query Match 14.2%; Score 71.5; DB 2; Length 887;
Best Local Similarity 26.1%; Pred. No. 71;
Matches 23; Conservative 23; Mismatches 37; Indels 5; Gaps 2;
Qy 18 AAAPVLSGAKVAPVQAVPVALESAAEAGAGTLANPLGTNPLKLLSS-LGIPVNHLEGSQ 76
Db 512 AGELTYGIIPGLEKELEAAAEARDSSGAGSMVQEVVTPDNIAHVVSRTWTGIPVDKMLEGQR 571
Qy 77 KCVA-----ELGPPQAVGAVKALKALLGAL 100
Db 572 EKLLRMEDELAKSVVGGQGEAVQAVSKAV 599

RESULT 21
C83368
probable MFS transporter PA2214 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: C83368
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim
; Jory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: C83368
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-440 <STO>
A:Cross-references: GB:AE004648; GB:AE004091; NID:g9948237; PIDN:AAG05602.1; GSPDB:GN00
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2214

Query Match 14.1%; Score 71; DB 2; Length 440;
Best Local Similarity 32.7%; Pred. No. 39;
Matches 36; Conservative 7; Mismatches 45; Indels 22; Gaps 5;
Qy 3 LAALLGLCVALLSCSSAAAFVLSGAKPV--AQPVVALESAAEAGAGTLANPLGTIN----- 55
Db 333 LAGNLGVIATVAFCAASGFV--SVQPLFWTLPTGYLSGAAAASGIALINSLNGLGFFVAP 390
Qy 56 PLKLLSSLGIPVNHLEGSQKCVAEIGPQAVGAVKALK-LIGALTVPFG 104
Db 391 NLKTLNE-----SQFADPRAGNFALAAVGLLGALLARLKTSG 428
```

Best Local Similarity 27.6%; Pred. No.: 83;  
Matches 24; Conservative 17; Mismatches 41; Indels 5; Gaps 2;

QY 18 AAFLVGSAPVQAQVAALSAEAGAGTLANPLGTINPLKLLSSLGIPVNHLSGSK 77  
||| : : : : : : : : : : : : : : : : : : : : : : : :  
Db 465 AAVLYKGAIQEVSALAIAKLEAKSDNY-MLTETVGPNIAEVSVRWTPVTRLDQNEKK 523

QY 78 CVAELG----PQAVCAVKALKALLGAL 100  
||| : : : : : : : : : : : : : : : : : : : : : : : :  
Db 524 RIISLADKLHERVVGGDEAVRAAAAI 550

RESULT 24  
C86178  
hypothetical protein [imported] - Arabidopsis thaliana  
C;Species:Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C;Accession: C86178  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alons  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.  
ansen, N.P.; Hughes, B.; Huizar, L.  
Nature 409, 816-820, 2000  
A;Authors: Hunter, J.B.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Marzial  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; UID:21016719; PMID:11130712  
A;Accession: C86178  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-431 <STO>  
A;Cross-references: GB:AE005172; NID:52494120; PIDN:AAB80629.1; GSPDB:GN00141  
C;Genetics:  
A;Map position: 1

Query Match 13.9%; Score 70; DB 2; Length 431;  
Best Local Similarity 28.2%; Pred. No. 47;  
Matches 35; Conservative 15; Mismatches 50; Indels 24; Gaps 4;

QY 3 LAALGLCVALSCTSAAFVGSAPVQAQVAALSAEAGAGTLANPLGTINPLKLL 61  
||| : : : : : : : : : : : : : : : : : : : : : : : :  
Db 122 LGNLGGVLLLTTPPKISFLVFSAULLSLQLVWVLSUSKEESFGLPRATSSVLESVKQI 161

QY 62 SSL-----GIPVNHLSGSKC---VAELGPQAVGA VKAL-KALLG 98  
||| : : : : : : : : : : : : : : : : : : : : : : : :  
Db 182 SNLKEAIQADETSQPLIWAVVSIAMVP LLSGSFVCYTQVINLPDSVIGMSKVIGQLMLL 241

QY 99 ALTIV 102  
|||  
Db 242 CLTV 245

RESULT 25  
T46147  
zinc finger protein - Arabidopsis thaliana  
N;Alternate names: protein TJAS.80  
C;Species:Arabidopsis thaliana (mouse-ear cress)  
C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
C;Accession: T46147  
R;Bloeker, H.; Meves, H.W.; Mayer, K.F.X.; Lemcke, K.; Schueller, C.; Quetier, F.; Sa  
submitted to the Protein Sequence Database, December 1999  
A;Reference number: ZZ3024  
A;Accession: T46147  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-452 <BLO>  
A;Cross-references: EMBL:AL132979  
A;Experimental source: cultivar Columbia; BAC clone TJAS5  
C;Genetics:  
A;Map position: 3  
A;Introns: 44/1; 176/2; 243/1

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Superfamily: Synchocystis hypothetical protein slr1977

Query Match 13.8%; Score 69.5; DB 2; Length 244;  
Best Local Similarity 31.2%; Pred. No. 30;  
Matches 30; Conservative 11; Mismatches 30; Indels 25; Gaps 5;

QY 17 SAAPFLVGSAPVAPQVAALSAEAGAGTIANPL-----GTIANPLKLLSLGI 66  
DB 160 NTAFLAFGEAKQI--PVAILRVSD---TMTODLPDINGVFTTEQALQPLAKALLR 213  
QY 67 PV--NHLIEGSKQKVAELGPOAVGAVKALKALCAL 100  
DB 214 FLAAGHLIQLSLKACQQLTAIA-----QSLSGAL 242

RESULT 28  
G75454  
hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C/Accession: G75454  
R/Heidelberg, J.F.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
Chardson, D.; Emolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, R.  
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of the chromosomes of the choleta pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A/Accession: D82179  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-477 <HEI>  
A/Cross-references: GB:AE004237; GB:AE003852; NID:G9656107; PIDN:AAF94751.1; GSPDB:GN001  
A/Experimental source: serogroup O1; strain N16961; biotype E1 Tor  
C:Genetics:  
A/Gene: VC1597  
A/Map position: 1  
C:Superfamily: multidrug-efflux transporter

Query Match 13.8%; Score 69.5; DB 2; Length 396;  
Best Local Similarity 32.1%; Pred. No. 48;  
Matches 34; Conservative 18; Mismatches 33; Indels 21; Gaps 5;

QY 3 LAALIGLCVALSCSAAAFVGSAPVAPQVAALSAEAGAGTIANPLKLL 61  
DB 38 LAGSAGWPGAAQSAAPAVRTLPASRSVPALPAAAPAPSGPLPASLLSSLSLNL-- 95  
QY 62 SLSLGPVNHIEGSKQK--VAELGPOAV-----GAVKALKALLG 98  
DB 96 ----IPI-----GSRLEVLGVAPDLAGTFCNPGALLAVRGVG 132

RESULT 29  
C70582  
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C/Accession: C70582  
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;  
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A/Reference number: A70500; MUID:198295987; PMID:9634230  
A/Accession: C70582  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-423 <COL>  
A/Cross-references: GB:Z95210; GB:AL123456; NID:G3261757; PIDN:CAB08513.1; PID:e315217;  
A/Experimental source: strain H37RV  
C:Genetics:  
A/Gene: PPE

A>Note: T3A5.80

Query Match 13.9%; Score 70; DB 2; Length 452;  
Best Local Similarity 26.5%; Pred. No. 50;  
Matches 28; Conservative 21; Mismatches 37; Indels 18; Gaps 5;

QY 9 LCVALLSCSAAAFVGS-----AKPVAQFVAALSAEAGAGTIANPLKLL 60  
DB 303 ISLGLSTSHGSSFL--GSRNFAQFAMSAATLIQAAQMGAAASGSLHGLIG-----I 354  
QY 61 LSSLGIPVNHIEGSKQKVAELGPOAVGAVKALKALCALVFG 104  
DB 355 VSSTSTSIDAIVPHGLGLGLPCGGSSGLKEL--WMGNSSVFG 396

RESULT 26  
D82179  
probable multidrug transporter VC1597 [imported] - Vibrio cholerae (strain N16961 serogroup  
C:Species: Vibrio cholerae  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 06-Jan-2003  
C/Accession: D82179  
R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
Chardson, D.; Emolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, R.  
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of the chromosomes of the choleta pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A/Accession: D82179  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-477 <HEI>  
A/Cross-references: GB:AE004237; GB:AE003852; NID:G9656107; PIDN:AAF94751.1; GSPDB:GN001  
A/Experimental source: serogroup O1; strain N16961; biotype E1 Tor  
C:Genetics:  
A/Gene: VC1597  
A/Map position: 1  
C:Superfamily: multidrug-efflux transporter

Query Match 13.9%; Score 70; DB 2; Length 477;  
Best Local Similarity 28.1%; Pred. No. 53;  
Matches 38; Conservative 16; Mismatches 39; Indels 42; Gaps 8;

QY 2 KLAALIGLCVALSCSAAAFVGSAPVAPQVAALSAEAGAGTIANPLKLL 47  
DB 332 KFLPVGMSLS--SCAFVAMSLSPAPVVSATIAL---CGLGLGVMTSTQVLTLAGK 387  
QY 48 ANPLGTINPLKLLSLGIPVNHIEGSKQK--VAELGPOAV-----QKCVAILGPOA 86  
DB 388 AN--LGRITAMASLSRSLGASVGTAFPGTILYSLPGLSPNSGLQIAAL--POSEILHAFQ 445  
QY 87 VG--AVKALKALCAL 100  
DB 446 IGFAVAALLALCAL 460

RESULT 27  
S75653  
hypothetical protein slr1977 - Synchocystis sp. (strain PCC 6803)  
C:Species: Synchocystis sp.  
C/Variety: PCC 6803  
C/Accession: S75653  
C/Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synchocystis  
A/Reference number: S74322; MUID:97061201; PMID:8905231  
A/Accession: S75653  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-244 <KAN>

```
Query Match      13.8%; Score 69.5; DB 2; Length 423;
Best Local Similarity 31.8%; Pred. No. 52;
Matches 28; Conservative 12; Mismatches 37; Indels 11; Gaps 3;

QY 16 SGAALFLVGSAPV-----AQPVALESAAEAGAGTANPLGTL-----NPLKLLSSLG 65
DB 158 SAAASALPFTPVQGTGPAGPAAATAATQAGAGAVADAQATLAQLPPGILSDILSALA 217
QY 66 IPNVHLIEGSKQCVAEELGPQAVGAVKAL 93
DB 218 ANADPLTSLGLGIASTLNPO-VGSAQPI 244

RESULT 30
AD0782
A:Title: L-serine ammonia-lyase (EC 4.3.1.17) [imported] - Salmonella enterica subsp. en
C:Species: Salmonella enterica subsp. enterica serovar typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AD0782
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, F.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Garra, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; UID:21534947; PMID:11677608
A:Accession: AD0782
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-455 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD02578.1; PID:g16503436; GSPDB:GN00176
C:Genetics:
A:Gene: STY2430
C:Superfamily: microbial L-serine dehydratase
C:Keywords: ammonia-lyase; carbon-nitrogen lyase; carbon-oxygen lyase; hydro-lyase

Query Match      13.8%; Score 69.5; DB 2; Length 455;
Best Local Similarity 29.0%; Pred. No. 56;
Matches 31; Conservative 17; Mismatches 26; Indels 33; Gaps 7;

QY 4 AALLGL---C---VALSCSAA---AFLVGSAPVQAPVALESAAEAGAGTANPLGTL 54
DB 329 ASILGSEVGCGGIGVACNSAAGLAELMGAS--VEQTLSAAEIAWEHHLGTCDFLG-- 384
QY 55 NPLKLLSSIGIPNVHLIEGSKQCVAEELGPQAVGAVKALKALGALT 101
DB 385 -----GQVQIP-----CIER---NAISAVKAINAATMAMS 411

RESULT 31
AF3312
A:Title: hypothetical protein BMEI0484 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AF3312
R:DeIvecchio, V.; Kapral, V.; Redkar, R.J.; Patra, G.; Muter, C.; Los, T.; Ivanova,
.; Mazur, M.; Gotsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252; PMID:11756688
A:Accession: AF3312
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-170 <KUR>
A:Cross-references: GB:AB008917; PIDN:AAL51665.1; PID:g17982396; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI0484
A:Map position: 1

Query Match      13.7%; Score 69; DB 2; Length 170;
Best Local Similarity 29.8%; Pred. No. 23;

Matches 31; Conservative 14; Mismatches 45; Indels 14; Gaps 5;

QY 6 LLGLCVALLSCSAAAFVLSAKPVAQPVALESAAEAGAGTANPLGTLNPLKLLSSL- 64
DB 2 LAGLAAACNSTESALDIQSSNKTGQ--AATTAPSNAPVATPAPQRAATLPGKLIHPIV 59
QY 65 GIPVN-----HLIEGSKQCVAE-ELG-----PQAVGAVKALKAL 97
DB 60 GAPVNVVTFPLTHRMNDDAKAMGIELAGNNDPSAAAYVIKGYFSVL 103

RESULT 32
D75393
A:Title: serine proteinase, subtilase family - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: D75393
R:White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; UID:20036896; PMID:10567266
A:Accession: D75393
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-627 <WHI>
A:Cross-references: GB:AE001990; GB:AE00513; NID:G6459214; PIDN:AAF11026.1; PID:g6459
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRI459
A:Map position: 1

Query Match      13.7%; Score 69; DB 2; Length 627;
Best Local Similarity 31.1%; Pred. No. 86;
Matches 32; Conservative 14; Mismatches 39; Indels 18; Gaps 5;

QY 2 KLAALIGLGVALLSCSAAAFVLSAKPVAQPVALESAAEAGAGTANPLGTLNPLKLL 61
DB 6 KULTLLGALALUSGASQAS--AGSLSP-----TLQKARAGDQT---PIGVIVRFNVAN 54
QY 62 SSIGIPNVHLIEGS-QKCVAEELGP-----QAVGAVKALKAL 97
DB 55 TAQGRALFKNLQNLNQIAKLGPSAGFLKQAVNSQKATQLWL 97

RESULT 33
T36551
A:Title: probable ATP-dependent proteinase ATP-binding chain - Streptomyces coelicolor (fragmen
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Feb-2001
C:Accession: T36551
R:Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z21609
A:Accession: T36551
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-853 <YUR>
A:Cross-references: EMBL:AL049754; PIDN:CAB42048.1; GSPDB:GN00070; SCOEDB:SCH10.39c
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: clpB; SCOEDB:SCH10.39c
C:Superfamily: endopeptidase Clp ATP-binding chain
C:Keywords: ATP; molecular chaperone; nucleotide binding; P-loop
F:195-202/Region: nucleotide-binding motif A (P-loop)
F:276-281/Region: nucleotide-binding motif B
F:597-604/Region: nucleotide-binding motif A (P-loop)
F:665-670/Region: nucleotide-binding motif B
F:201/Binding site: ATP (lys) #status predicted
F:603/Binding site: ATP (lys) #status predicted

Query Match      13.7%; Score 69; DB 2; Length 853;
```

Best Local Similarity 27.8%; Pred. No. 1.2e+02;  
Matches 25; Conservative 21; Mismatches 38; Indels 6; Gaps 4;

QY 17 SAAFLVGSAPVQPV-AALSAAGAGTTLA-NPLGTLNPLKLLSLGIPVNHLEIG 74  
DB 487 TASKLLYGEIPDERLEAAAEAEVEEVARDTWVKEVGADDIADVAVSWGTPAGRLLEG 546

QY 75 -SOKCVA---ELGFPQAVGAVKALKALIGAL 100  
DB 547 ETQKLMEDELKRLIGQAVRAVSDAV 576

RESULT 34  
AE0074  
probable adhesin YPO0599 [imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001  
C:Accession: AE0074  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Ill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AE0001; MUID:21470413; PMID:11586360  
A:Accession: AE0074  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-3295 <KUR>  
A:Cross-references: GB:AL590842; PIDN:CAC89456.1; PID:g15978692; GSPDB:GN00175  
C:Genetics:  
A:Gene: YPO0599

Query Match 13.7%; Score 69; DB 2; Length 3295;  
Best Local Similarity 27.6%; Pred. No. 4.6e+02;  
Matches 32; Conservative 15; Mismatches 43; Indels 26; Gaps 5;

QY 4 AALGLC-----VALSCSSAAAFVVG-----SAKPAQVVAALSAAGAGTIL 48  
DB 2948 AAVGGLAGNMGALT-GASAPYLAGVVKOSTGDPNPAANTMAHVLGAVTAVASGNALA 3006

QY 49 NPLGTINP-----LKLLSLGIPVNHLEIGSKVQVAILGPOAVGAVKALKALIGALT 101  
DB 3007 GAAGAATAELMAPTIISALGDKNTLTGEGQAVSALTSLAAG-----LAGGLT 3055

RESULT 35  
F84336  
hypothetical protein Vng1852h [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: F84336  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo, Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.F.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
A:Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; MUID:20504483; PMID:11016950  
A:Accession: F84336  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-277 <STO>  
A:Cross-references: GB:AE004437; NID:g10581298; PIDN:AAG20058.1; GSPDB:GN00138  
C:Genetics:  
A:Gene: VNG1852H

Query Match 13.6%; Score 68.5; DB 2; Length 277;  
Best Local Similarity 27.5%; Pred. No. 42;  
Matches 33; Conservative 14; Mismatches 33; Indels 33; Gaps 5;

QY 3 LAALLGLCVALS-----FLVG-SAKPAQPV-AALSAAGAG 44  
DB 71 VAPVFGSPVIVATTAAAGATWLAVRFGARFQVLLAGVGAAVAAVITAALARGWTAGT 130

QY 45 GTLANPLGTINPLKLLSLGIPVNHLEIGSKVQVAILGPOAVGAVKALKALIGALT 104  
DB 131 LTLAWFLGAVGAALVLT-----GWRLLARIRPTDTRAV-----GVAGVLA 175

RESULT 36  
AE0575  
ferric enterobactin transport protein FepD STY0636 [imported] - Salmonella enterica sub  
C:Species: Salmonella enterica subsp. enterica serovar Typhi  
A:Note: this species has also been called Salmonella typhi  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C:Accession: AB0575  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, th, T.; Conneron, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero  
A:Reference number: AB0502; MUID:21534947; PMID:11677608  
A:Accession: AB0575  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-335 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAD05068.1; PID:g16501843; GSPDB:GN00176  
C:Genetics:  
A:Gene: STY0636  
C:Superfamily: vitamin B12 transport protein btuC

Query Match 13.6%; Score 68.5; DB 2; Length 335;  
Best Local Similarity 19.5%; Pred. No. 51;  
Matches 33; Conservative 22; Mismatches 45; Indels 69; Gaps 6;

QY 3 LAALLGLCVALS-----SAAFLVGS-----KPAQV----- 32  
DB 41 LEAFTGCGSADCTIVLDARLPTLAGLLAGALGALMOTLTENPLADPGILGVANG 100

QY 33 -----VAALSAABAGAGTILANPL-----GTNPLKLLSLGIPV 68  
DB 101 ASPAIVLGAALGFSTPLEQFMFSGALIASLIIVFTSGSGGQLSPVRLTLA--GVAL 158

QY 69 NHLEIGSKVQVAILGPOA-----VGAKVAKALKALIGALT 104  
DB 159 GAVLEGLTSGIALLNPEVVDQURFWQAGSLDRSQTLLKVALAPVVIAG 207

RESULT 37  
T18343  
P-glycoprotein - Sauroleishmania tarentolae  
C:Species: Sauroleishmania tarentolae  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
C:Accession: T18343  
R:Legare, D.; Hetttema, E.; Ouellette, M.  
Mol. Biochem. Parasitol. 68, 81-91, 1994  
A:Title: The P-glycoprotein-related gene family in Leishmania.  
A:Reference number: Z18879; MUID:95198776; PMID:7891750  
A:Accession: T18343  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1724 <LEG>  
A:Cross-references: EMBL:L29485; NID:g460312; PID:g460313; PIDN:AAA65541.1  
C:Superfamily: human multidrug resistance protein CMOAT2; ATP-binding cassette homology

Query Match 13.6%; Score 68.5; DB 2; Length 1724;  
Best Local Similarity 29.2%; Pred. No. 2.7e+02;  
Matches 28; Conservative 10; Mismatches 35; Indels 23; Gaps 4;

QY 25 SAKPAQVAAA---LESAAABAGAG---TLANPLGTINP-----LKLLSLSLGI 66  
DB 722 SADPAMEPLTTERTEGHGDAASPEPLSSPAITINSEVVFNEDSYTLNTKLTLDVNL 781

QY 67 PVNHLEIGSKVQVAILGPOAVGAVKALKALIGALT 102

Db 782 RVPR-----GRLTVLGLPTGSGKSTLLDALIGALAV 812

## RESULT 38

S16296  
ferric enterobactin transport protein fepD - Escherichia coli (strain K-12)  
N/Alternate names: ferric enterobactin permease fepD  
C/Species: Escherichia coli  
C/Date: 04-Dec-1992 #sequence revision 04-Dec-1992 #text\_change 01-Mar-2002  
C/Accession: S16296; S16305; D64792; S14841  
R/Shea, C.M.; McIntosh, M.A.  
Mol. Microbiol. 5, 1415-1428, 1991  
A/Title: Nucleotide sequence and genetic organization of the ferric enterobactin transport protein fepD  
A/Reference number: S16295; MUID:92157868; PMID:1838574  
A/Accession: S16296  
A/Molecule type: DNA  
A/Residues: 1-334 <SHE>  
A/Cross-references: EMBL:X57471; NID:G41429; PIDN:CAA40707.1; PID:G41430  
R/Chenault, S.S.; Earhart, C.F.  
Mol. Microbiol. 5, 1405-1413, 1991  
A/Title: Organization of genes encoding membrane proteins of the Escherichia coli ferric enterobactin transport system  
A/Reference number: S16305; MUID:92157867; PMID:1787794  
A/Accession: S16305  
A/Molecule type: DNA  
A/Residues: 1-232, 'AL', 235-334 <CHE>  
A/Cross-references: EMBL:X59402; NID:G41433; PIDN:CAA42043.1; PID:G41434  
R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, D.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A/Title: The complete genome sequence of Escherichia coli K-12.  
A/Reference number: A64720; MUID:97426617; PMID:9278503  
A/Accession: D64792  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-334 <BLAT>  
A/Cross-references: GB:AE000164; GB:U00096; NID:G1786800; PIDN:AC73691.1; PID:G1786805;  
A/Experimental source: strain K-12, substrain MG1655  
C/Genetics:  
A/Gene: fepD  
A/Map position: 14 min  
C/Superfamily: vitamin B12 transport protein btuC  
C/Keywords: transmembrane protein  
F/16-32/Domain: transmembrane #status predicted <TM1>  
F/65-81/Domain: transmembrane #status predicted <TM2>  
F/94-110/Domain: transmembrane #status predicted <TM3>  
F/121-137/Domain: transmembrane #status predicted <TM4>  
F/156-172/Domain: transmembrane #status predicted <TM5>  
F/198-214/Domain: transmembrane #status predicted <TM6>  
F/240-256/Domain: transmembrane #status predicted <TM7>  
F/284-300/Domain: transmembrane #status predicted <TM8>  
F/310-326/Domain: transmembrane #status predicted <TM9>

Query Match 13.5%; Score 68; DB 2; Length 334;  
Best Local Similarity 22.1%; Pred. No. 56;  
Matches 34; Conservative 19; Mismatches 43; Indels 58; Gaps 6;  
Qy 3 LAALLGLCVALS-----SAAFLVGS-----KPAQ----- 32  
Db 41 LEAFSGTCQSDCTIVLDARLPRTLAGLGGALGALMOTLRNPLADPGLLVNAG 100  
Qy 33 -----VAALSSAAAGAGTLANPL-----GTLNPKLLSSLGIPV 68  
Db 101 ASFAIVLGAALFGYSSAQQLNAFAGALVSLIVAFVTSQGGQLSPVRLTLA--GVAL 158

Qy 69 NHIIEGSKVAGELGPQAVGAVKALKALLGALT 102  
Db 159 AAVLEGLTSGIALLNPVDVYDLRFWQA--GSLDI 190

## RESULT 39

A85558  
ferric enterobactin (enterochelin) transport [imported] - Escherichia coli (strain O157:  
C/Species: Escherichia coli

C/Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 17-May-2002  
C/Accession: A85558  
R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, M.W.; Davis, N.W.; Lim, A.; Dimantanta, E.; Potamousis, K.; Apodaca, L.; Grotbeck, E.J.; Nature 409, 529-533, 2001  
A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A/Reference number: A85480; MUID:21074935; PMID:11206551  
A/Accession: A85558  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-334 <STO>  
A/Cross-references: GB:AE005174; NID:G12513480; PIDN:AAG54925.1; GSPDB:GN00145; UWGP:Z  
A/Experimental source: strain O157:H7, substrain EDL933  
C/Genetics:  
A/Gene: fepD  
C/Superfamily: vitamin B12 transport protein btuC

Query Match 13.5%; Score 68; DB 2; Length 334;  
Best Local Similarity 22.1%; Pred. No. 56;  
Matches 34; Conservative 19; Mismatches 43; Indels 58; Gaps 6;  
Qy 3 LAALLGLCVALS-----SAAFLVGS-----KPAQ----- 32  
Db 41 LEAFSGTCQSDCTIVLDARLPRTLAGLGGALGALMOTLRNPLADPGLLVNAG 100  
Qy 33 -----VAALSSAAAGAGTLANPL-----GTLNPKLLSSLGIPV 68  
Db 101 ASFAIVLGAALFGYSSAQQLNAFAGALVSLIVAFVTSQGGQLSPVRLTLA--GVAL 158

Qy 69 NHIIEGSKVAGELGPQAVGAVKALKALLGALT 102  
Db 159 AAVLEGLTSGIALLNPVDVYDLRFWQA--GSLDI 190

## RESULT 40

E90707  
ferric enterobactin (enterochelin) transport E90629 [imported] - Escherichia coli (st  
C/Species: Escherichia coli  
C/Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 17-May-2002  
C/Accession: E90707  
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.  
DNA Res. 8, 11-22, 2001  
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and ge  
A/Reference number: A99629; MUID:21156231; PMID:11258796  
A/Accession: E90707  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-334 <HAY>  
A/Cross-references: GB:BA000007; PIDN:BA34052.1; PID:G13360087; GSPDB:GN00154  
A/Experimental source: strain O157:H7, substrain RIMD 0509952  
C/Genetics:  
A/Gene: E90629  
C/Superfamily: vitamin B12 transport protein btuC

Query Match 13.5%; Score 68; DB 2; Length 334;  
Best Local Similarity 22.1%; Pred. No. 56;  
Matches 34; Conservative 19; Mismatches 43; Indels 58; Gaps 6;  
Qy 3 LAALLGLCVALS-----SAAFLVGS-----KPAQ----- 32  
Db 41 LEAFSGTCQSDCTIVLDARLPRTLAGLGGALGALMOTLRNPLADPGLLVNAG 100  
Qy 33 -----VAALSSAAAGAGTLANPL-----GTLNPKLLSSLGIPV 68  
Db 101 ASFAIVLGAALFGYSSAQQLNAFAGALVSLIVAFVTSQGGQLSPVRLTLA--GVAL 158

Qy 69 NHIIEGSKVAGELGPQAVGAVKALKALLGALT 102  
Db 159 AAVLEGLTSGIALLNPVDVYDLRFWQA--GSLDI 190

Search completed: April 5, 2004, 14:36:01



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 5, 2004, 14:11:13 ; Search time 17 Seconds

(without alignments)

318.547 Million cell updates/sec

Title: US-09-997-428-408

Perfect score: 502

Sequence: 1 MKLALLGLCVLSCSSAAA.....QAVGAVKALKALGALTVEG 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt\_42:\*

Sequence ID, Comparison

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	497	99.0	104	1	UGR2 HUMAN	Q96qr1	homo sapien
2	250	49.8	104	1	UGR2 MOUSE	Q920d7	mus musculus
3	160	31.9	93	1	UGR1 HUMAN	Q96p11	homo sapien
4	135	26.9	139	1	UGR1 MOUSE	Q920h1	mus musculus
5	83	16.5	1327	1	TKN1 HUMAN	Q95271	homo sapien
6	77	15.3	335	1	TRD2 STRCO	Q924w9	streptomyces
7	76	15.1	732	1	COPA HELPE	Q32619	helicobacter
8	75	14.9	362	1	MDCL HUMAN	Q9hik6	homo sapien
9	75	14.9	362	1	MDCL MOUSE	Q90670	mus musculus
10	74	14.7	779	1	SNIL MOUSE	Q8d357	wiggleswort
11	73.5	14.6	626	1	DXS WIGBR	Q9rlu5	rattus norv
12	73.5	14.6	776	1	SNIL RAT	P74941	thermus the
13	73.5	14.6	882	1	SYA THETH	Q8uc60	agrobacteri
14	72	14.3	397	1	SUCC AGRTS	Q84733	mus musculus
15	71	14.1	428	1	FXE2 MOUSE	P35890	rhizobium s
16	71	14.1	480	1	DNAH RHIME	Q8xuz7	raistonia s
17	69	13.7	124	1	RL7 RALSO	Q8g4t9	bifidobacte
18	68.5	13.6	377	1	PROB BIFLO	P32876	escherichia
19	68	13.5	334	1	PEPD ECOLI	Q920g3	mus musculus
20	68	13.5	569	1	SIUF MOUSE	Q94508	homo sapien
21	67.5	13.4	2035	1	GCP MYCLE	P37969	mycobacteri
22	67	13.3	351	1	FUMC PSESM	Q885v0	pseudomonas
23	67	13.3	464	1	TAZR RAT	P34978	rattus norv
24	66.5	13.2	341	1	FXDI HUMAN	Q16676	homo sapien
25	66.5	13.2	465	1	NLTD BRAOL	Q43304	brassica ol
26	66	13.1	118	1	MP70 MYCKA	Q49614	mycobacteri
27	66	13.1	121	1	CLRC IDEDE	P60000	ideonella d
28	66	13.1	239	1	PTB MOUSE	P17225	mus musculus
29	66	13.1	527	1	LEUI PROMM	Q7tuv5	prochloroco
30	66	13.1	540	1	LEUI PROMM	P70581	rattus norv
31	66	13.1	585	1	NUPI RAT	Q59782	rhodococcus
32	65.5	13.0	257	1	EUTC ROHER	Q98kn9	rhizobium 1
33	65.5	13.0	336	1	COBT RHIL0		

RESULT 1

See over

ALIGNMENTS

34	65.5	13.0	491	1	Y084 MYCTU	O53209	mycobacteri
35	65	12.9	344	1	TRD2 RALSO	Q8xe00	raistonia s
36	65	12.9	399	1	YM96 SYNEL	Q8dgm0	synecococc
37	65	12.9	462	1	A2AC HUMAN	P18825	homo sapien
38	65	12.9	593	1	REGG ECOLI	P24230	escherichia
39	65	12.9	4377	1	ANK3 HUMAN	Q12955	homo sapien
40	64.5	12.8	184	1	CYCM BRAUJA	P30323	bradyrhizob
41	64.5	12.8	320	1	Y678 METJA	Q58091	methanococc
42	64.5	12.8	705	1	ELS CHICK	P07916	gallus gall
43	64.5	12.8	950	1	NUO3 PSEAF	Q910j6	pseudomonas
44	64.5	12.8	1023	1	RT11 ACTPL	P55128	actinobacil
45	64.5	12.8	1023	1	RT12 ACTPL	P55129	actinobacil
46	64.5	12.8	1159	1	RPOC PORCN	O33431	porphyromon
47	64.5	12.8	1608	1	HLXA SERMA	P15320	serattia ma
48	64.5	12.8	2541	1	TLN1 HUMAN	Q9490	homo sapien
49	64	12.7	246	1	HIS4 METMA	Q8pws2	methanosarc
50	64	12.7	337	1	THI4 ASPOR	Q9uuz9	aspergillus
51	64	12.7	426	1	Y200 ANASP	Q82098	anabaena sp
52	64	12.7	487	1	DNAH AGRT5	Q8uih1	agrobacteri
53	64	12.7	655	1	YAGF ECOLI	P77596	escherichia
54	64	12.7	693	1	REGG ECO57	Q8xd86	escherichia
55	63.5	12.6	82	1	C551 PSEME	P00102	pseudomonas
56	63.5	12.6	130	1	F14A HUMAN	Q942x8	homo sapien
57	63.5	12.6	130	1	RL7 MYCLE	P30763	mycobacteri
58	63.5	12.6	230	1	PLGH XANCP	Q8p9b9	xanthomonas
59	63.5	12.6	345	1	TRPD AERPE	Q9y8t2	aeropyrum p
60	63.5	12.6	344	1	YGAY ECO57	Q8x4v6	escherichia
61	63.5	12.6	394	1	YGAY ECOLI	P76628	escherichia
62	63.5	12.6	397	1	Y435 PROMA	P59918	prochloroco
63	63.5	12.6	407	1	AROC MYCLE	Q9ccs6	mycobacteri
64	63.5	12.6	432	1	YF10 MYCTU	P71789	mycobacteri
65	63.5	12.6	541	1	Y4MM RHISN	P55572	rhizobium s
66	63.5	12.6	548	1	MERA PSEFL	Q51772	pseudomonas
67	63.5	12.6	1048	1	P100 HCMVA	P08318	human cytom
68	63	12.5	434	1	YU18 MYCTU	P21500	mycobacteri
69	63	12.5	435	1	YU21 MYCTU	O53268	mycobacteri
70	63	12.5	517	1	6PGD CANAL	O13287	candida alb
71	63	12.5	932	1	PMS1 HUMAN	P54277	homo sapien
72	63	12.5	1062	1	NAL2 HUMAN	Q9nx02	homo sapien
73	62.5	12.5	267	1	THIG DEIRA	Q9ryv1	deinococcus
74	62.5	12.5	361	1	COBT MYCTU	Q10396	mycobacteri
75	62.5	12.5	511	1	NADP CAUCR	Q944c3	caulobacter
76	62.5	12.5	560	1	EFS MOUSE	Q64355	mus musculus
77	62.5	12.5	760	1	CO2 MOUSE	P21180	mus musculus
78	62	12.4	128	1	RL7 CORGL	Q8nt28	corynebacte
79	62	12.4	199	1	IL11 RAT	Q99mf5	rattus norv
80	62	12.4	242	1	HIS4 HALN1	Q9hm14	halobacteri
81	62	12.4	257	1	MOAB KLEAE	P54795	klebsiella
82	62	12.4	299	1	G3P DICDI	Q94469	dictyosteli
83	62	12.4	382	1	CHSD PHANI	P71884	mycobacteri
84	62	12.4	388	1	PCAB PSEPU	O22045	pharbitis n
85	62	12.4	446	1	MURF STTUA	P32427	pseudomonas
86	62	12.4	460	1	SMP COTUA	O33804	streptomyce
87	62	12.4	620	1	REP2 HUMAN	Q92154	coturnix co
88	62	12.4	660	1	REP2 HUMAN	Q8nf8	homo sapien
89	62	12.4	660	1	REP2 HUMAN	P51010	xanthobacte
90	62	12.4	678	1	TKT XANFL	P14380	xenopus lae
91	62	12.4	775	1	YTX1 XENLA	P20105	drosophila
92	62	12.4	829	1	E74B DROME	P11536	drosophila
93	62	12.4	883	1	GCSP AGRT5	Q8ufd6	agrobacteri
94	61.5	12.3	124	1	RL7 BRUNE	P41106	brucella me
95	61.5	12.3	217	1	RNH2 AGRT5	Q8uhg2	agrobacteri
96	61.5	12.3	306	1	DDL HAFIN	P44495	haemophilus
97	61.5	12.3	318	1	KLPF XANCP	P23354	xanthomonas
98	61.5	12.3	344	1	NAGK HUMAN	Q9uJ70	homo sapien
99	61.5	12.3	355	1	TPPD AZOBR	P26924	azospirillu
100	61.5	12.3	367	1	NK61 HUMAN	P78426	homo sapien



```
UGR2_HUMAN
ID UGR2_HUMAN STANDARD; PRT; 104 AA.
AC Q96P01;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uteroglobin-related protein 2 precursor (Cytokine HIN-1) (High in
DE normal-1) (Secretoglobins family 3A member 1).
GN SCGB3A1 OR UGRP2 OR HIN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21396515; PubMed=11481438;
RA Krop I.E., Sgroi D., Porter D.A., Lunetta K.L., LeVangie R., Seth P.,
RA Kaelin C.M., Rhei E., Rosenberg M., Schnitt S., Marks J.R., Pagon Z.,
RA Belina D., Razumovic J., Polyak K.;
RT "HIN-1, a putative cytokine highly expressed in normal but not
RT cancerous mammary epithelial cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9796-9801(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21539178; PubMed=11682631;
RA Niimi T., Keck-Waggoner C.L., Popescu N.C., Zhou Y., Levitt R.C.,
RA Kimura S.;
RT "UGRP1, a uteroglobin/clara cell secretory protein-related protein, is
RT a novel lung-enriched downstream target gene for the T/EBP/NKX2.1
RT homeodomain transcription factor.";
RL Mol. Endocrinol. 15:2021-2036(2001).
CC -!- FUNCTION: Potential growth inhibitory cytokine.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Highly expressed in breast tissues. Absent in
CC breast cancer cell lines.
CC -!- SIMILARITY: Belongs to the uteroglobin family. UGRP subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
EMBL: AY040564; AAK82942.1; .
EMBL: AF313458; AL26217.1; .
Genew; HGNC:18384; SCGB3A1.
DR MIN; 606500; .
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0005125; P:cytokine activity; NAS.
DR GO; GO:0030308; P:negative regulation of cell growth; NAS.
DR GO; GO:0042127; P:regulation of cell proliferation; NAS.
KW Cytokine; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 104 UTEROGLOBIN-RELATED PROTEIN 2.
FT CONFLICT 19 19 R -> A (IN REF. 2).
SQ SEQUENCE 104 AA; 10185 MW; 1083673C8FAE8015 CRC64;
Query Match 99.0%; Score 497; DB 1; Length 104;
Best Local Similarity 99.0%; Pred. No. 7.2e-36;
Matches 103; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MKLALLGLCVALLSCSSAAFLVGSAPVQVPAALSAEAAAGAGTLANPLGTINPLKLL 60
DB 1 MKLALLGLCVALLSCSSAAFLVGSAPVQVPAALSAEAAAGAGTLANPLGTINPLKLL 60
QY 61 LSSLGIPVNHLEIGSKCVKVAELGPQAVGAVKALKALLGALTVEG 104
DB 61 LSSLGIPVNHLEIGSKCVKVAELGPQAVGAVKALKALLGALTVEG 104
```

RESULT 2

```
UGR2_MOUSE
ID UGR2_MOUSE STANDARD; PRT; 104 AA.
AC Q920D7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uteroglobin-related protein 2 precursor (Cytokine HIN-1) (High in
DE normal-1) (Secretoglobins family 3A member 1).
GN SCGB3A1 OR UGRP2 OR HIN1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21396515; PubMed=11481438;
RA Krop I.E., Sgroi D., Porter D.A., Lunetta K.L., LeVangie R., Seth P.,
RA Kaelin C.M., Rhei E., Rosenberg M., Schnitt S., Marks J.R., Pagon Z.,
RA Belina D., Razumovic J., Polyak K.;
RT "HIN-1, a putative cytokine highly expressed in normal but not
RT cancerous mammary epithelial cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9796-9801(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21539178; PubMed=11682631;
RA Niimi T., Keck-Waggoner C.L., Popescu N.C., Zhou Y., Levitt R.C.,
RA Kimura S.;
RT "UGRP1, a uteroglobin/clara cell secretory protein-related protein, is
RT a novel lung-enriched downstream target gene for the T/EBP/NKX2.1
RT homeodomain transcription factor.";
RL Mol. Endocrinol. 15:2021-2036(2001).
CC -!- FUNCTION: Potential growth inhibitory cytokine.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- SIMILARITY: Belongs to the uteroglobin family. UGRP subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
EMBL: AF313456; AAL26216.1; .
DR MGD; MGI:1915912; Scgb3a1.
KW Cytokine; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 104 UTEROGLOBIN-RELATED PROTEIN 2.
SQ SEQUENCE 104 AA; 10591 MW; D62F0B601FB57A6D CRC64;
Query Match 49.8%; Score 250; DB 1; Length 104;
Best Local Similarity 57.0%; Pred. No. 8.2e-16;
Matches 61; Conservative 14; Mismatches 26; Indels 6; Gaps 4;
QY 1 MKLITFVLCVALLSDSGVAFVMDSLAKPAVEPVAAAPAAEAAGVAVPSPLSHLAIL 60
DB 1 MKLITFVLCVALLSDSGVAFVMDSLAKPAVEPVAAAPAAEAAGVAVPSPLSHLAIL 60
QY 58 KLLSLGIPVNHLEIGSKCVKVAELGPQAVGAVKALKALLGALTVEG 104
DB 61 RFLASMGIPLDPIEGSKCVKVAELGPQAVGAVKALKALLGALTVEG 104
RESULT 3
UGR1_HUMAN
ID UGR1_HUMAN STANDARD; PRT; 93 AA.
AC Q96P01;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Uteroglobin-related protein 1 precursor (Secretoglobins family 3A
DE member 2).
GN SCGB3A2 OR UGRP1.
```



31-JUL-1998; 98US-0094983P.  
01-OCT-1998; 98US-0102686P.  
11-DEC-1998; 98US-0112129P.  
(INCY-) INCYTE PHARM INC.  
Lal P, Tang YT, Gorgone GA, Corley NC, Guegler KJ, Baughn WR;  
Akerblom IE, Au-Young J, Yue H, Patterson C, Reddy R, Hillman JL;  
Bandman O;  
WPI; 2000-160673/14.  
N-PSDB; AAZ98173.  
New human signal peptide-containing proteins useful in treatment,  
prevention and diagnosis of e.g. cancer, inflammation and cardiovascular  
disease.  
Claim 1; Page 206; 327pp; English.  
AAZ98109 to AAZ98242 encode AAY87224 to AAY87357 which represent the  
human signal peptide-containing proteins HSP-1 to HSP-134. HSPs have  
anticancer, anti-inflammatory, antimicrobial, neurotropic, hepatotropic,  
neuroprotective, cardiovascular and antiasthmatic activities, and can be  
used in gene therapy. HSPs can be used to treat or prevent disorders  
associated with decreased activity or function of HSP. Antagonists of  
HSP are used to treat or prevent disorders associated with increased  
activity or function of HSP. Such diseases include cell proliferation  
(including cancer), inflammation, cardiovascular, neurological,  
reproductive or developmental disorders, (e.g. arteriosclerosis,  
cirrhosis, psoriasis, acquired immune deficiency syndrome, anaemia,  
asthma, Crohn's disease, Alzheimer's, Parkinson's or Huntington's  
diseases, schizophrenia, ovulatory defects, muscular dystrophy). HSP  
nucleic acids can be used for the recombinant production of HSP, for  
detecting HSP in standard hybridisation and amplification assays (for  
diagnosis and monitoring), in gene therapy, as antisense, triplex-forming  
or ribozyme therapeutics, for detecting related sequences or genetic  
variations, and for chromosomal mapping. HSP are also used to raise  
specific antibodies (Ab) and to screen for agonists and antagonists  
(potential therapeutic agents). Ab are used to diagnose, or monitor, HSP  
-related diseases (in usual immunoassays), as therapeutic antagonists, in  
competitive drug screens, and for purification of HSP from natural  
sources  
Sequence 104 AA;  
Query Match 100.0%; Score 502; DB 3; Length 104;  
Best Local Similarity 100.0%; Pred. No. 1.4e-48;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKLALLGCVALLSCSSAAAFVLSAKPVQVVALESAAAGAGTIANPIGTINPLKLL 60  
Db 1 MKLALLGCVALLSCSSAAAFVLSAKPVQVVALESAAAGAGTIANPIGTINPLKLL 60  
QY 61 LSLGLPVNHLIEGSKVAVLGPQVAVKALKALLGALTVEG 104  
Db 61 LSLGLPVNHLIEGSKVAVLGPQVAVKALKALLGALTVEG 104  
RESULT 4  
AAB65280  
ID AAB65280 standard; protein; 104 AA.  
XX AAB65280;  
AC AAB65280;  
XX 02-APR-2001 (first entry)  
XX Human PRO1245 (UNQ629) protein sequence SEQ ID NO:408.  
XX Human; secreted and transmembrane protein; PRO; cytosstatic; cell death;  
XX cancer; chromosomal mapping; Gene mapping; tissue typing;  
XX diagnostic assay.

OS Homo sapiens.  
XX WO2000073454-A1.  
XX 07-DEC-2000.  
XX 30-MAR-2000; 2000WO-US008439.  
XX 02-JUN-1999; 99WO-US012252.  
XX 23-JUN-1999; 99US-0141037P.  
PR 07-JUL-1999; 99US-0143048P.  
PR 20-JUL-1999; 99US-0144758P.  
PR 26-JUL-1999; 99US-0145698P.  
PR 28-JUL-1999; 99US-0146222P.  
PR 17-AUG-1999; 99US-0149396P.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 08-OCT-1999; 99US-0158663P.  
PR 30-NOV-1999; 99WO-US028313.  
PR 01-DEC-1999; 99WO-US028301.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030911.  
PR 05-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
(GETH) GENENTECH INC.  
PA Askenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;  
PI Ferrara N, Fong S, Gerber H, Gritsen ME, Goddard A, Godowski PJ;  
PI Grimaldi CJ, Gurney AL, Kijavini LJ, Napier MA, Pan J, Paoni NF;  
PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;  
PI Zhang Z;  
XX WPI; 2001-032160/04.  
DR N-PSDB; AAF44249.  
XX PRO polynucleotides used to produce polypeptides used to target bioactive  
PT molecules such as toxins, radiolabels or antibodies, to specific cells,  
PT to cause targeted cell death.  
XX Claim 12; Fig 290; 935pp; English.  
XX The present invention describes human secreted and transmembrane PRO  
CC proteins. The PRO proteins have cytostatic activity. The PRO proteins can  
CC be used for targeted delivery of bioactive molecules, such as toxins,  
CC radiolabels or antibodies, that cause cell death. PRO nucleotide  
CC sequences, and their fragments, can be used as hybridisation probes, in  
CC chromosomal and gene mapping, and in the generation of anti-sense RNA and  
CC DNA. They may also be used to produce transgenic animals which are used  
CC to develop and screen therapeutically useful reagents. The PRO nucleotide  
CC and protein sequence can be used for tissue typing and in treating  
CC cancer. Anti-PRO antibodies can be used in diagnostic assays. AAF44270 to  
CC AAF44470 represent PCR primers and hybridisation probes used in the  
CC isolation of human PRO sequences. AAF44087 to AAF44269 and AAB65154 to  
CC AAB65300 represent human PRO polynucleotide and protein sequences given  
CC in the exemplification of the present invention  
XX Sequence 104 AA;  
SQ Query Match 100.0%; Score 502; DB 4; Length 104;  
Best Local Similarity 100.0%; Pred. No. 1.4e-48;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKLALLGCVALLSCSSAAAFVLSAKPVQVVALESAAAGAGTIANPIGTINPLKLL 60

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 5, 2004, 14:09:33 ; Search time 54 Seconds

(without alignments)

544.166 Million cell updates/sec

Title: US-09-997-428-408

Perfect score: 502

Sequence: 1 MKLAALGLCVLSCSSAAA.....QAVGAVKALKALLGALTTFVG 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A\_Geneseq\_29Jan04.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	502	100.0	104	3	AAV66757 Membrane-
2	502	100.0	104	3	AAV44458 Human lun
3	502	100.0	104	3	AAV87288 Human sig
4	502	100.0	104	4	AB65280 Human PRO
5	502	100.0	104	5	ABU8141 Human PRO
6	502	100.0	104	6	ABU58095 Human PRO
7	502	100.0	104	6	ABU59173 Novel hum
8	502	100.0	104	6	ABU82685 Human sec
9	502	100.0	104	6	AAO19895 Human ute
10	502	100.0	104	6	ABU60604 Human sec
11	502	100.0	104	6	ABU13986 Human PRO
12	502	100.0	104	6	ABU72571 Novel hum
13	502	100.0	104	6	ABU59320 Human sec
14	502	100.0	104	6	ABO26017 Human PRO
15	502	100.0	104	6	ABU59026 Human sec
16	502	100.0	104	6	ABU92404 Novel hum
17	502	100.0	104	6	ABU59469 Human PRO
18	502	100.0	104	6	ABU92235 Novel hum
19	502	100.0	104	6	ABU10941 Human PRO
20	502	100.0	104	6	ABU81693 Novel hum
21	502	100.0	104	6	ABU88632 Human sec
22	502	100.0	104	6	ABO34146 Human PRO
23	502	100.0	104	6	ADA37919 Human sec
24	502	100.0	104	6	ADA21605 Human sec
25	502	100.0	104	6	ADA10392 Human sec

26	502	100.0	104	6	ADA17936	Human PRO
27	502	100.0	104	6	ADA28044	Human sec
28	502	100.0	104	6	ADA94624	Human sec
29	502	100.0	104	6	ADA38849	Human sec
30	502	100.0	104	6	ADA92370	Human sec
31	502	100.0	104	7	ABO53332	Human sec
32	502	100.0	104	7	ADA22531	Human sec
33	502	100.0	104	7	ABO22602	Human sec
34	502	100.0	104	7	ADA06697	Human sec
35	502	100.0	104	7	ADA39390	Human PRO
36	502	100.0	104	7	ADB96416	Human PRO
37	502	100.0	104	7	ADC57888	Human PRO
38	502	100.0	104	7	ADC55432	Human PRO
39	502	100.0	104	7	ADC12119	Human PRO
40	502	100.0	104	7	ADC65641	Human PRO
41	502	100.0	104	7	ADC07596	Human sec
42	502	100.0	104	7	ADC11586	Human sec
43	502	100.0	104	7	ADC14708	Novel hum
44	502	100.0	104	7	ADD08240	Novel hum
45	502	100.0	104	7	ADC82065	Novel hum
46	502	100.0	104	7	ADD07707	Novel hum
47	502	100.0	104	7	ADC82598	Human PRO
48	502	100.0	104	7	ADD08778	Novel hum
49	502	100.0	104	7	ADD07027	Novel hum
50	502	100.0	104	7	ADC83274	Human PRO
51	502	100.0	104	7	ADD55381	Human PRO
52	502	100.0	104	7	ADD56339	Human PRO
53	502	100.0	104	7	ADD54777	Human PRO
54	502	100.0	104	7	ADE26931	Novel hum
55	502	100.0	104	7	ADE26398	Novel hum
56	498	99.2	104	2	AAW75868	Novel hum
57	250	49.8	104	2	AAO19893	Murine ut
58	160	31.9	93	2	AAW62068	Human lun
59	160	31.9	93	3	AAV28334	LJ103 spe
60	160	31.9	93	3	AAV44456	Human lun
61	160	31.9	93	3	AAV87289	Human sig
62	160	31.9	93	4	AAU29145	Human PRO
63	160	31.9	93	6	ABU58521	Human PRO
64	160	31.9	93	6	ABU88069	Novel hum
65	160	31.9	93	6	ABU84384	Human sec
66	160	31.9	93	6	ABR6258	Human sec
67	160	31.9	93	6	ABR65648	Human sec
68	160	31.9	93	6	ABU99588	Human sec
69	160	31.9	93	6	ABU82827	Human PRO
70	160	31.9	93	6	ABU89948	Novel hum
71	160	31.9	93	6	ABR68197	Human sec
72	160	31.9	93	6	AAO19894	Human ute
73	160	31.9	93	6	ABU96250	Novel hum
74	160	31.9	93	6	ABU52681	Human sec
75	160	31.9	93	6	ABO08758	Human sec
76	160	31.9	93	6	ABO02810	Human sec
77	160	31.9	93	6	ABR74964	Human sec
78	160	31.9	93	6	ABR94726	Human sec
79	160	31.9	93	6	ABU85699	Human PRO
80	160	31.9	93	6	ABU98859	Novel hum
81	160	31.9	93	6	ABU98074	Novel hum
82	160	31.9	93	6	ABU91780	Novel hum
83	160	31.9	93	6	ABU9473	Human PRO
84	160	31.9	93	6	ABU86314	Human sec
85	160	31.9	93	6	ABU67527	Human sec
86	160	31.9	93	6	ABU80555	Human PRO
87	160	31.9	93	6	ABR9473	Human sec
88	160	31.9	93	6	ABR9863	Human sec
89	160	31.9	93	6	ABO16386	Human sec
90	160	31.9	93	6	ABR92286	Human sec
91	160	31.9	93	6	ABO18927	Human sec
92	160	31.9	93	6	ABR78348	Human sec
93	160	31.9	93	6	ABU85084	Novel hum
94	160	31.9	93	6	ABO00223	Novel hum
95	160	31.9	93	6	ABO11555	Human sec
96	160	31.9	93	6	ABO02200	Human sec
97	160	31.9	93	6	ABU88774	Novel hum
98	160	31.9	93	6	ABU83469	Human sec

Ab006270 Novel hum  
AbR59306 Human sec

99 160 31.9 93 6 ABO06270  
100 160 31.9 93 6 ABR59306

## ALIGNMENTS

## SEQUENCE COMPARISON

RESULT 1  
AAy66757  
ID AAY66757 standard; protein; 104 AA.

AC AAY66757;

XX 05-APR-2000 (first entry)

XX Membrane-bound protein PRO1245.

XX Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;  
KW pharmaceutical; receptor immunoadhesin; gene mapping.

XX Homo sapiens.

XX WO9963088-A2.

XX 09-DEC-1999.

XX 02-JUN-1999; 99WO-US012252.

XX 02-JUN-1998; 98US-0087607P.

XX 02-JUN-1998; 98US-0087609P.

XX 03-JUN-1998; 98US-0087759P.

XX 04-JUN-1998; 98US-0087827P.

XX 04-JUN-1998; 98US-0088021P.

XX 04-JUN-1998; 98US-0088025P.

XX 04-JUN-1998; 98US-0088028P.

XX 04-JUN-1998; 98US-0088029P.

XX 04-JUN-1998; 98US-0088030P.

XX 04-JUN-1998; 98US-0088033P.

XX 04-JUN-1998; 98US-0088326P.

XX 05-JUN-1998; 98US-0088167P.

XX 05-JUN-1998; 98US-0088202P.

XX 05-JUN-1998; 98US-0088212P.

XX 05-JUN-1998; 98US-0088217P.

XX 09-JUN-1998; 98US-0088655P.

XX 10-JUN-1998; 98US-0088722P.

XX 10-JUN-1998; 98US-0088730P.

XX 10-JUN-1998; 98US-0088734P.

XX 10-JUN-1998; 98US-0088738P.

XX 10-JUN-1998; 98US-0088740P.

XX 10-JUN-1998; 98US-0088741P.

XX 10-JUN-1998; 98US-0088742P.

XX 10-JUN-1998; 98US-0088810P.

XX 10-JUN-1998; 98US-0088811P.

XX 10-JUN-1998; 98US-0088824P.

XX 10-JUN-1998; 98US-0088825P.

XX 10-JUN-1998; 98US-0088826P.

PR 18-JUN-1998; 98US-0089908P.  
PR 19-JUN-1998; 98US-0089947P.  
PR 19-JUN-1998; 98US-0089948P.  
PR 19-JUN-1998; 98US-0089952P.  
PR 22-JUN-1998; 98US-0090246P.  
PR 22-JUN-1998; 98US-0090252P.  
PR 22-JUN-1998; 98US-0090254P.  
PR 22-JUN-1998; 98US-0090349P.  
PR 23-JUN-1998; 98US-0090355P.  
PR 23-JUN-1998; 98US-0090355P.  
PR 24-JUN-1998; 98US-0090429P.  
PR 24-JUN-1998; 98US-0090431P.  
PR 24-JUN-1998; 98US-0090435P.  
PR 24-JUN-1998; 98US-0090444P.  
PR 24-JUN-1998; 98US-0090445P.  
PR 24-JUN-1998; 98US-0090451P.  
PR 24-JUN-1998; 98US-0090472P.  
PR 24-JUN-1998; 98US-0090535P.  
PR 24-JUN-1998; 98US-0090538P.  
PR 24-JUN-1998; 98US-0090540P.  
PR 25-JUN-1998; 98US-0090557P.  
PR 25-JUN-1998; 98US-0090676P.  
PR 25-JUN-1998; 98US-0090678P.  
PR 25-JUN-1998; 98US-0090688P.  
PR 25-JUN-1998; 98US-0090690P.  
PR 25-JUN-1998; 98US-0090691P.  
PR 25-JUN-1998; 98US-0090694P.  
PR 25-JUN-1998; 98US-0090695P.  
PR 25-JUN-1998; 98US-0090696P.  
PR 26-JUN-1998; 98US-0090862P.  
PR 26-JUN-1998; 98US-0090863P.  
PR 01-JUL-1998; 98US-0091358P.  
PR 01-JUL-1998; 98US-0091360P.  
PR 02-JUL-1998; 98US-0091478P.  
PR 02-JUL-1998; 98US-0091486P.  
PR 02-JUL-1998; 98US-0091519P.  
PR 02-JUL-1998; 98US-0091544P.  
PR 02-JUL-1998; 98US-0091626P.  
PR 02-JUL-1998; 98US-0091628P.  
PR 02-JUL-1998; 98US-0091633P.  
PR 02-JUL-1998; 98US-0091646P.  
PR 07-JUL-1998; 98US-0091673P.  
PR 07-JUL-1998; 98US-0091978P.  
PR 07-JUL-1998; 98US-0091982P.  
PR 09-JUL-1998; 98US-0092182P.  
PR 10-JUL-1998; 98US-0092472P.  
PR 30-JUL-1998; 98US-0093339P.  
PR 04-AUG-1998; 98US-0094651P.  
PR 04-AUG-1998; 98US-0095282P.  
PR 04-AUG-1998; 98US-0095285P.  
PR 04-AUG-1998; 98US-0095301P.  
PR 04-AUG-1998; 98US-0095302P.  
PR 04-AUG-1998; 98US-0095318P.  
PR 04-AUG-1998; 98US-0095321P.  
PR 10-AUG-1998; 98US-0095325P.  
PR 10-AUG-1998; 98US-0095916P.  
PR 10-AUG-1998; 98US-0095929P.  
PR 11-AUG-1998; 98US-0096012P.  
PR 11-AUG-1998; 98US-0096143P.  
PR 11-AUG-1998; 98US-0096146P.  
PR 12-AUG-1998; 98US-0096329P.  
PR 17-AUG-1998; 98US-0096757P.  
PR 17-AUG-1998; 98US-0096766P.  
PR 17-AUG-1998; 98US-0096768P.  
PR 17-AUG-1998; 98US-0096773P.  
PR 17-AUG-1998; 98US-0096791P.  
PR 17-AUG-1998; 98US-0096867P.  
PR 17-AUG-1998; 98US-0096891P.  
PR 17-AUG-1998; 98US-0096894P.  
PR 17-AUG-1998; 98US-0096895P.  
PR 17-AUG-1998; 98US-0096897P.  
PR 18-AUG-1998; 98US-0096949P.  
PR 18-AUG-1998; 98US-0096950P.  
PR 18-AUG-1998; 98US-0096959P.